

State of California

---

# **The Status of the State's Decentralized Testing Program**

## ***Final Report of Findings and Recommendations***

---

**Prepared by**

**Policy Division  
State Personnel Board**



**January 2003**



**State of California**

**The Status of the State's  
Decentralized Testing Program**

***Final Report of Findings  
and Recommendations***



**Prepared by**

Policy Division  
Special Projects Unit  
January 2003

## **Table of Contents**

<b>Listing of Appendices</b>	<b>ii</b>
<b>Purpose of the Report</b>	<b>1</b>
<b>Background of the State's Decentralized Testing Program</b>	<b>3</b>
<b>Project Methodology</b>	<b>8</b>
Literature Review	10
Survey of State Government Testing Practices	10
Review of Quality Assurance Audit Findings	11
Review of Examination Appeals Data	11
Departmental Questionnaire	11
Departmental Interviews	12
Union and Employee Advocacy Group Interviews	14
Review of Decentralized Testing Processes	15
Cost Comparison	17
State Personnel Board Hearing	18
<b>Results and Findings of Data Collection</b>	<b>20</b>
Literature Review	20
Survey of State Government Testing Practices	26
Review of Quality Assurance Audit Findings	29
Review of Examination Appeals Data	34
Departmental Questionnaire	39
Departmental Interviews	52
Union Interviews	58
Employee Advocacy Group Interviews	61
Review of Decentralized Testing Processes	63
Cost Comparison	92
State Personnel Board Hearing	97
<b>Summary of Findings and Discussion</b>	<b>103</b>
<b>Recommendations</b>	<b>108</b>
<b>Appendices</b>	<b>115</b>

## Appendices

Appendix A	Participants in the <i>Survey of State Government Testing Practices</i>
Appendix B	<i>Decentralized Testing Questionnaire</i> and cover memo
Appendix C	Participants in the <i>Departmental Interviews</i>
Appendix D	Departmental Interview Script
Appendix E	Participants in the <i>Union and Employee Advocacy Group Interviews</i>
Appendix F	Union/Employee Advocacy Group Interview Script
Appendix G	<i>Departmental Selection Process Evaluation Criteria</i>
Appendix H	Groups Participating in the July 2002 Hearing before the State Personnel Board
Appendix I	Hearing Notice for July 2002 Hearing before the State Personnel Board (without attachment)
Appendix J	Scoring models comprising decentralized testing processes included in the <i>Review of Decentralized Testing Processes</i>
Appendix K	Cost comparison of centralized and decentralized examination processes
Appendix L	<i>Standards for Decentralized Departmental Examination Programs</i>

# **The Status of the State's Decentralized Testing Program**

## ***Final Report of Findings and Recommendations***

### **Purpose of Report**

Article VII of the California Constitution requires that all appointments and promotions in the State civil service be made on the basis of merit ascertained by competitive examination, making the State's selection system a merit-based system. The Constitution further requires that the State Personnel Board (SPB) oversee the State's selection system and enforce the civil service laws to ensure compliance with the merit principle. The State's selection system encompasses a myriad of selection functions, including recruitment and advertising, testing for the establishment of eligible lists, hiring interviews, background/reference checks, medical evaluation, drug testing, psychological screening, and civil service probation. Currently, the State's selection system is decentralized and provides for individual State departments and agencies, under the authority and oversight of the SPB, to administer their own selection processes, inclusive of initial recruitment and advertising efforts, testing for eligible list establishment, effecting hiring decisions, and administering civil service probationary periods.

One aspect of the State's selection system is the decentralized testing program, whereby State departments and agencies have delegated authority to conduct formal examination processes, including both the development and the administration of those processes. The decentralized testing program encompasses testing and selection decisions for civil service classifications, including Career Executive Assignment (CEA) positions, as well as testing and selection decisions made pursuant to Board-approved demonstration projects. Decentralized employment testing was originally implemented within State service on a pilot basis in 1981, and was fully implemented in the early 1990s.

Since its inception, no formal comprehensive review has been undertaken to evaluate the effectiveness of the decentralized testing program. To the extent that problematic testing procedures or practices may be utilized, the effectiveness of the decentralized testing program to identify qualified candidates for employment may be compromised. Without a systemic program review, the effectiveness of the



State's decentralized testing program is unknown.

In September 2001, the five-member State Personnel Board (Board) adopted a resolution calling for a comprehensive review and evaluation of the State's decentralized testing program, as well as a public hearing to solicit comments and input from program stakeholders. In addition, the resolution called for Board adoption of findings and recommendations identified through the review process. SPB staff has completed a comprehensive review of the State's decentralized testing program. The review focused on decentralized examination processes conducted for eligible list creation for civil service classifications. Examination processes conducted for demonstration projects and Career Executive Assignment appointments were not included in this review. In addition, while assessment components such as hiring interviews, psychological screening, background interviews, drug testing, and medical examinations are part of the merit selection process, these were not elements addressed in this study. This report documents the review process and outlines preliminary findings and recommendations in preparation for the public hearing.

## **Background of the State's Decentralized Testing Program**

A number of laws and rules establish the State's civil service and corresponding selection system. Article VII, Section 1(b), of the California Constitution requires that appointments and promotions in the State's civil service be made on the basis of merit ascertained by competitive examination. Government Code Section 18500(c)(2) requires that appointments and promotions be based upon merit and the fitness of the candidate as determined through competitive examination. Government Code Section 18930 requires that such competitive examinations fairly test and determine the qualifications, fitness, and ability of candidates to perform the duties of the classification to which they seek appointment. Government Code Sections 18950 and 18951 require that positions in the civil service be filled by the best-qualified candidates and that advancement in State service be based upon merit and ability. To this end, the State's civil service system is a merit-based system that utilizes formal employment testing procedures as one component of the overall hiring process to assess candidate qualifications for purposes of identifying those candidates best qualified for appointment.

Under current rules, the State's selection system provides for decentralized employment testing by State departments and agencies under the authority and oversight of the SPB. Decentralized employment testing was originally implemented within State service in 1981. Prior to the implementation of the decentralized testing program, employment testing for the State's civil service was conducted on a centralized basis by the SPB. Under provisions of the centralized testing program, the SPB was responsible for developing all examination processes, and the SPB shared administration of these centrally developed examination processes with State departments and agencies under provisions of the SPB-implemented Delegated Testing Program. The Delegated Testing Program provided for the delegation of examination administration activities only to State departments and agencies on an examination-by-examination basis, with strict SPB pre-approval and rigid procedural oversight requirements. The Delegated Testing Program allowed for departments and agencies to administer under the close scrutiny of the SPB, SPB-developed examination processes.

In August 1979, the Little Hoover Commission published a report on the Personnel Management System in State service. A major concern discussed in the report was the fact that the centralized employment testing program was not meeting the needs

of State service. The centralized system, including delegated testing provisions in place, was noted as being complex, rigid, and protracted. In follow-up documentation, a 1980 report sanctioned by the Joint Legislative Audit Committee, and conducted by the consulting firm of Ralph Anderson and Associates, stated that through the decentralization of the testing process from the SPB to State departments, a more effective use of resources and significant improvements in testing could be achieved. As noted in the report, decentralized testing activities included both examination development and administration activities, and decentralization of both activities was advocated in the report. As a result of findings in the Little Hoover Commission report, recommendations in the Ralph Anderson and Associates report, and a subsequent study conducted by a task force of SPB staff, a pilot decentralized testing program was conducted.

Under the pilot decentralized testing program, examination planning, development, and administration activities were delegated to State departments and agencies. Participating departments were responsible for all phases of the examination process resulting in the establishment of eligible lists, including examination planning, development of individual selection instruments, examination bulletin preparation, application review, examination administration, scoring, documentation, and records retention. The pilot program called for the SPB to establish standards and guidelines by which departments and agencies would be required to conduct their examination activities. The SPB was also called upon to develop and provide training to departmental staff engaged in conducting examination activities, as well as provide consultative assistance to departments to ensure the development and administration of merit-based examination processes. The SPB also developed processes to monitor and review the results of the pilot program to ensure that decentralized examination processes were merit-based and resulted in the identification of individuals eligible for appointment on the basis of job-related assessment. The pilot decentralized testing program began on April 1, 1981, and included the participation of the Franchise Tax Board and the Departments of Motor Vehicles, General Services, and the Youth Authority.

The pilot program implementation methodology included the selection of departmental participants that were representative of the State's geographic diversity and widely diverse job classifications. Those departments selected also had to express a willingness to pursue the development and use of fair, merit-based examination processes. Based upon departmental staff expertise and corresponding training needs, the SPB developed and conducted training for departmental staff to ensure that staff were adequately educated on a variety of assessment-related topics. Topics addressed in the training included the principles of the State's merit system; the importance of job-relatedness in selection activities; strengths and weaknesses of various testing procedures (e.g., written examinations,



interviews); examination statistics and their uses in assessment activities; and, the implications and requirements of the federal *Uniform Guidelines on Employee Selection Procedures* (29 CFR 1607). The training was intended to provide departmental staff with the expertise and resources needed to effectively conduct required testing activities under the parameters of decentralized testing. Further, the implementation methodology called for considerable consultation and oversight of departmental testing activities by SPB staff. During the pilot program, for all decentralized testing processes, SPB staff reviewed examination planning efforts and the examination bulletin prior to its release. In addition to pre-review activities, SPB staff also were available to provide consultation and guidance to departmental staff on a variety of issues and matters related to the development and administration of decentralized testing processes. The final element in the implementation methodology entailed the SPB conducting post-administration audits of the decentralized testing processes. The post-administration audits consisted of a structured review of all aspects of the testing process, including examination planning, test development, recruitment and publicity efforts, application review activities, the appropriateness of testing instruments used, examination security efforts, and the certification of eligible candidates.

In reports to the Legislature in December 1981, and February 1982, documenting the results of the pilot decentralized testing program, the SPB concluded that the pilot program had been successful. Of primary note was the fact that the selected departments could administer examination processes on a decentralized basis significantly faster with no increase in costs than equivalent examination processes could be administered through the State's centralized testing program. Based on the success of the pilot program, the SPB's reports recommended that decentralization be expanded to other State departments. The SPB's recommendation called for the utilization of a structured implementation methodology, modeled after that which was used in the pilot program, as well as on-going monitoring of departmental decentralized testing efforts. The expanded implementation methodology called for the SPB to provide extensive training and consultation to the departments engaged in decentralized testing activities. The methodology also required departments to complete a probationary period during which the department's decentralized testing processes would be thoroughly evaluated by SPB staff. In addition, the SPB reports noted that not all examination processes were suitable for decentralization and, therefore, recommended that testing for large, servicewide classifications remain centralized within the SPB.

As a result of the pilot program and the SPB's 1981 and 1982 reports to the Legislature, plans to expand the decentralized testing program were developed by the SPB in early 1982. Under expansion plans, it was intended that decentralized testing authority would be extended on a voluntary basis to a minimum of six

departments/agencies annually until such time as most departments/agencies were participants in the decentralized testing program. Noted exceptions to participation in the program included small departments and those departments and agencies that declined to participate. It was anticipated that a structured implementation methodology as outlined in the pilot program reports would be used to expand the decentralized testing program. Further, program expansion plans indicated that SPB staff who had previously been engaged in conducting centralized testing for classifications that would eventually be decentralized should be transitioned to those departments which would assume decentralized testing authority. It was anticipated that existing resources would simply be reallocated from the SPB to decentralized testing departments; no reduction nor expansion in resources was anticipated.

In September 1982, Assembly Bill 3332, which had been passed by the Legislature, was signed into law. AB 3332 formally institutionalized the wide-scale implementation of decentralized testing to departments that were interested and qualified to conduct testing processes. AB 3332 called for the SPB to “authorize or assess the ability of appointing powers to design, announce, or administer designated examinations for the establishment of employment lists...” by January 1, 1987. Per AB 3332 and implementation activities already underway, the SPB continued expansion of the decentralized testing program throughout the State, adhering to the recommended implementation methodology discussed above, including the one-year probationary period, to ensure the appropriate training and performance of departmental participants in the program. In addition, the SPB notified departments that upon completion of the one-year probationary period, staff resources commensurate with the reduction in centralized examining services and decentralized testing oversight would be deleted from the SPB's budget and made available for allocation to departments. Throughout expansion efforts, the SPB retained testing authority for servicewide classifications.

In 1985, implementation of the decentralized testing program was accelerated from the January 1, 1987, date stipulated in AB 3332 to June 30, 1985, with the objective being to implement decentralized testing in most State departments and agencies. The SPB hastened implementation of the decentralized testing program in the State's civil service, with the exception of testing for large, servicewide classifications, which was maintained on a centralized basis under the SPB. Due to the acceleration of implementation and the volume of departments newly enrolled in the decentralized testing program, the SPB was unable to provide the degree of training and oversight to departments as had been recommended in the implementation methodology. In addition, budget constraints precluded the reallocation of SPB staff resources to the decentralized departments. Thus, departments were afforded decentralized testing authority without the benefit of adequate staffing and staff training and preparation, as well as SPB staff oversight

and guidance. Further, the one-year probationary period was in essence eliminated from the implementation process.

In July 1987, subsequent to the accelerated 1985 decentralized testing implementation efforts, the SPB authorized departments with decentralized testing authority to conduct promotional testing for servicewide classifications, with the SPB retaining authority for open testing for servicewide classifications. Finally, in the early 1990s, the SPB suffered funding cuts, necessitating that open testing for servicewide classifications, which had yet to be decentralized, be decentralized to State departments.

In the late 1980s and early 1990s, as the State's testing system was moving toward decentralization, many departments remained interested in the SPB conducting their testing. As departments adjusted to their increased authority and accountability for conducting much of the State's testing, some departments requested that they be allowed to contract with the SPB to conduct their testing on a reimbursable basis. Thus, for many years in the 1990s, the SPB provided testing expertise and testing services on a reimbursable basis, conducting examination processes for departments on an as-needed, contractual basis.

In 1999, after several years of decentralized testing and reimbursable SPB testing services, some funding for centralized testing was restored to the SPB, enabling the SPB to conduct minimal testing for servicewide classifications. At present, the SPB conducts approximately 11 servicewide examination processes annually, which equates to approximately 0.3% of the State's total annual examining for the civil service. The remaining testing for the civil service is conducted on a decentralized basis by individual State departments and agencies. The SPB is available on a reimbursable, contractual basis to assist departments in the development and administration of testing processes; however, the SPB's role in reimbursable testing activities is minimal compared to the amount of testing conducted for State service. Throughout the fluctuations and demands of decentralization, the SPB has remained focused on sound, innovative assessment, providing testing expertise to the State's testing program.

## Project Methodology

Since its inception, no formal evaluation of the State's decentralized testing program has been conducted. Absent a systemic program review, the effectiveness of decentralized testing is unknown. To determine the effectiveness of the decentralized testing program within State service, a comprehensive evaluation was conducted by SPB staff. The goals of the evaluation were to identify any deficiencies or problems with the program, as well as to identify the positive attributes of the program. The review process focused on examination processes conducted for civil service classifications, excluding CEA examinations and demonstration projects.

Departmental testing processes and practices were evaluated to determine the extent to which legal and professional standards and principles were emulated. The evaluation included a review of the decentralized testing program in terms of the degree to which departmental testing practices adhered to State law and regulations pertaining to employment testing practices and recognized legal and professional testing standards and principles, including the *Uniform Guidelines*, the Civil Rights Acts of 1964 and 1991, the Americans with Disabilities Act, the American Psychological Association, Division 14, *Principles for the Validation and Use of Personnel Selection Procedures* (1987), and the American Psychological Association *Standards for Educational and Psychological Testing* (1999). In addition, program performance was compared to findings and recommendations in the professional literature. A growing body of professional literature (see for example *The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 85 Years of Research Findings* (Schmidt and Hunter, 1998) and *Beyond Employment Interview Validity: A Comprehensive Narrative Review of Recent Research and Trends over Time* (Posthuma, Morgeson, and Campion, 2002) demonstrates that the job-relatedness and predictive value of personnel selection practices are maximized when highly structured and standardized assessment techniques are used.

The evaluation methodology consisted of the following data collection elements:

1. A review of the professional literature to identify research findings on decentralized testing practices within the human resources discipline

2. A survey of other state governments with merit-based civil service systems to determine the extent of decentralized testing practices
3. A review of audit findings by the SPB's Quality Assurance unit to identify common issues and/or patterns of quality deficiencies or efficiencies in decentralized testing processes currently developed and administered by State departments
4. A review of examination appeal data to identify trends and/or common issues in appeals resulting from decentralized testing processes currently developed and administered by State departments
5. The dissemination of a comprehensive questionnaire to all State departments engaged in decentralized testing activities to solicit data on testing practices, staff development, and resources required to implement decentralized testing
6. Interviews with personnel officers and/or examination managers in a sampling of small, medium, and large departments to solicit data and comments on departmental decentralized testing practices and the status of the decentralized testing program
7. Interviews with union representatives to solicit data and comments on the State's decentralized testing program
8. Interviews with employee advocacy group representatives to solicit data and comments on the State's decentralized testing program
9. An audit of a sample of decentralized testing processes to determine the quality, soundness, job-relatedness, and fairness of the processes
10. A cost comparison of decentralized testing efforts versus centralized testing efforts

Each of the above-noted data collection elements is described in detail as follows.

## **Literature Review**

A review of the professional literature was conducted to identify research findings and/or published studies addressing decentralized testing practices within the human resources field. The literature review focused on data collection points in the following areas:

- The degree to which decentralized testing programs are utilized by other public-sector organizations
- The extent to which decentralized testing programs have been studied
- The extent to which the qualitative aspects of decentralized testing programs have been measured
- The degree to which decentralized testing efforts are monitored or audited by a central testing authority

The literature review was conducted in December 2001, and January 2002.

## **Survey of State Government Testing Practices**

A review of the testing practices of other state governments which utilize a merit-based selection model was conducted. It should be noted that the criterion of a merit-based selection model did not require that the state utilize a civil service system. The review was designed to identify the following information:

- The extent to which decentralized testing is utilized
- For states utilizing decentralized testing, the extent to which the decentralized testing programs are monitored or reviewed for quality and effectiveness
- For states not utilizing decentralized testing programs, the extent to which the states are investigating the use of and/or the implementation of decentralized testing programs.

A total of 14 states were identified for survey purposes, and the surveys were conducted in early 2002. A listing of the states surveyed is included in *Appendix A*.

## **Review of Quality Assurance Audit Findings**

A review of audit findings documented by the SPB's Quality Assurance unit was conducted to identify any trends, deficiencies, and/or efficiencies with decentralized testing processes. The review focused on the following issues:

- The degree to which departmental decentralized testing efforts result in sound, quality assessment of candidate qualifications
- The presence of any common deficiencies in decentralized testing processes
- The expertise of departmental examination staff conducting decentralized testing processes

This review was conducted in December 2001.

## **Review of Examination Appeals Data**

A review of appeals records for the fiscal years of 1998/1999, 1999/2000, and 2000/2001 was conducted to identify any trends or appeals activities resulting directly from decentralized testing processes. The review consisted exclusively of examination appeals; records of other types of appeals were not reviewed. Issues considered in the review process included the following:

- The total number of appeals filed per fiscal year
- The number of appeals granted per fiscal year
- The number of appeals filed against individual departments
- The volume of appeals filed against specific testing processes
- The bases of the appeals
- The outcome of the appeals

This review was conducted in January 2002.

## **Departmental Questionnaire**

A comprehensive questionnaire was disseminated to the 89 State departments currently engaged in decentralized testing. The questionnaire elicited data regarding each department's testing activities for the fiscal years of 1998/1999, 1999/2000,



and 2000/2001. Data collected for each fiscal year included the following:

- The number of examination processes administered
- The types of selection instruments utilized (e.g., written examinations, interviews, performance tests)
- The examination bases (e.g., open, promotional, open, non-promotional)
- The number of candidates processed
- The job analytic data upon which the examination processes were based (e.g., job analysis, classification specifications, subject matter expert input)

The questionnaire also solicited general demographic data about departmental examination staff. The demographic data documented the size of the department's examination staff, the number and levels of employees engaged in examination activities, and the training and expertise of the examination staff. Additionally, the questionnaire asked departments to indicate their use of a variety of SPB products and services. The questionnaire identified the products and services currently being used by departments and their examination staff, the frequency with which those products and services are accessed, and the products and services that are of the greatest value to departments and their examination staff.

The questionnaire was distributed February 15, 2002, with an initial requested return date of March 8, 2002. Extensions to the return date deadline were offered and afforded to departments, and a final deadline of April 29, 2002, was established. A copy of the *Decentralized Testing Questionnaire* and a copy of the cover memo which accompanied it are included in *Appendix B*.

## **Departmental Interviews**

A sampling of small, medium, and large departments engaged in decentralized testing was interviewed by project staff. The interviews were designed to collect additional qualitative and quantitative data that could not effectively be collected via the questionnaire described previously. In selecting the sample departments, it was determined that a sample size of 15, representing 17 percent of the departments engaged in decentralized testing, would be of adequate size to ensure appropriate representation. The following criteria were taken into consideration in determining the department sample:



- Achieving an equal distribution of small, medium, and large departments in the sample
- Ensuring adequate representation of departments utilizing a variety of classifications with varied types and levels of work performed, as well as a variety of union representation
- Inclusion of departments utilizing the State's largest hiring classifications
- Inclusion of departments with classifications for which examination processes consistently result in very large candidate groups
- Inclusion of departments with diverse and/or specialized classifications for which recruitment and/or retention might present challenges
- Inclusion of departments utilizing peace officer and safety personnel

Consideration of these criteria resulted in the identification of five large, five medium, and five small departments, as well as the SPB, for inclusion in the interview process. The sample of 16 departments represented 18 percent of the departments engaged in decentralized testing. A listing of the departments participating in the interview process is included in *Appendix C*.

The interview was designed to collect data regarding departmental testing practices and solicit comments on the effectiveness of the decentralized testing program. A standardized interview, comprised of 10 questions, some of which included multiple parts, was used. The interview elicited data and information regarding each department's specific testing practices, as well as perceptions of the decentralized testing program in general. Questions included in the interview addressed a variety of topics related to departmental administration of decentralized testing, including:

- The job analytic data upon which examination processes developed and administered by the departments are based
- The extent to which subject matter experts are utilized in the development and administration of examination processes
- The extent to which departments utilize available training offered through the SPB's Selection Analyst Training Program

- The extent to which departments apply the concepts and principles presented in the SPB's Selection Analyst Training Program classes to their departmental examination activities
- The extent to which training is provided to hiring supervisors/managers in the conducting of hiring interviews and the selection of candidates to fill vacancies

In addition to the data noted above, the interview also elicited data regarding the effectiveness of decentralized testing from the departments' perspective. The following questions were used to assess the departments' perceptions of the quality and effectiveness of the decentralized testing program:

- *What is currently working with the decentralized testing program?*
- *What currently is not working with the decentralized testing program?*
- *How could the decentralized testing program be improved?*
- *What is your perception of the Board's [SPB's] role in the decentralized testing program?*
- *What could the Board [SPB] do, or do more of, to improve the decentralized testing program?*

The interviews were held during March and April 2002. A copy of the complete interview questions is included in *Appendix D*.

### **Union and Employee Advocacy Group Interviews**

The 12 unions representing the State's 21 bargaining units and seven employee advocacy groups were contacted regarding this project. Representatives from these groups were invited on behalf of their respective memberships to provide data and information regarding the status of the decentralized testing program.

Following the same format as the interviews held with the sample of departments, the interview was designed to collect data and comments regarding the effectiveness of the decentralized testing program. A standardized interview, comprised of three questions, one of which included multiple parts, was used. The interview elicited data regarding each group's perception of the effectiveness of

decentralized testing processes. Similar to the questions posed to the sampled departments, the following questions were used to assess the groups' perceptions of the quality and effectiveness of the decentralized testing program:

- *What is currently working with the State's decentralized testing program?*
- *What currently is not working with the decentralized testing program?*
- *What parameters or provisions of the decentralized testing program, if any, should be discontinued?*
- *What could the SPB do, or do more of, to improve the decentralized testing program?*

The interviews were held during March and April 2002. A listing of the unions and employee advocacy groups that participated in the interview process is included in *Appendix E*. A copy of the interview questions is included in *Appendix F*.

## **Review of Decentralized Testing Processes**

In order to obtain objective data pertaining to the effectiveness of decentralized testing, a formal audit/evaluation of a sampling of decentralized testing processes was completed. This review used structured evaluation criteria across a multitude of variables to determine the quality, soundness, and fairness of testing processes developed and administered on a decentralized basis. In selecting the testing processes for review, it was determined that approximately three to five processes from each of the 16 departments included in the interview sample would be evaluated. It was further determined that similar numbers of processes would be reviewed for small, medium, and large departments to ensure appropriate representation by department size.

The following criteria were taken into consideration in identifying the testing processes to be included in the evaluation:

- Recency of completed processes
- A mix of open and promotional processes
- Processes with varying numbers of candidates

- Processes conducted for entry-level, journey-level, supervisory, and managerial classifications
- Processes comprised of multiple selection instruments, as well as those comprised of only a single instrument
- Processes conducted for specialized classifications, as well as for more general classifications
- Processes conducted for department-specific classifications, as well as for servicewide classifications

Based upon these criteria, 86 testing processes were reviewed. The 86 processes included 29 processes conducted by large departments, 29 conducted by medium departments, and 28 conducted by small departments.

The review process utilized standardized evaluation criteria to evaluate the quality, soundness, and fairness of each of the processes, including:

- The job analytic data upon which the testing processes were developed
- The extent to which subject matter experts were utilized in the development and administration of the testing processes
- The appropriateness of the selection instruments used
- The appropriateness of the scoring model used
- The soundness of the selection instruments developed
- The standardization of administration of the selection instruments
- The presence of adverse impact in each selection instrument used in the testing process
- The presence of adverse impact at the bottom line of the testing process

The reviews were conducted during March and April 2002. A copy of the evaluation criteria is included in *Appendix G*.

## Cost Comparison

A cost comparison between centralized and decentralized testing was conducted. The purpose of the cost comparison was to determine whether decentralized testing efforts are any more costly than comparable centralized testing efforts. In conducting the preliminary research for the cost comparison and designing the methodology for the collection and analysis of the data, project staff determined that departments conducting examination processes did not routinely track actual costs incurred in the development and administration of examination processes. The lack of readily available cost data made planning the cost comparison process somewhat difficult. To ensure that an accurate comparison could be made between centralized and decentralized testing, and that appropriate cost data were collected, the cost comparison focused on a single servicewide job classification. Project staff worked closely with departmental staff to collect cost data reflective of known costs, as well as estimated and projected costs.

The large, servicewide classification of Office Technician (Typing) was selected to be the subject of the cost comparison. The SPB had conducted a centralized open examination process for the classification of Office Technician (Typing) in the Spring of 2002. Data from the SPB's 2002 servicewide open examination process for Office Technician (Typing) represented cost data for centralized testing. Data for decentralized testing processes were collected from a sample of departmental examination processes conducted for the Office Technician (Typing) classification. The following criteria were taken into consideration in identifying the departmental participants in the data collection effort to ensure the comparability between the centralized and decentralized cost data:

- The recency of the departmental examination processes
- The size of the applicant/candidate groups assessed in the departmental examination processes
- The use of a written examination as either the sole selection instrument or one of multiple selection instruments in the departmental examination processes
- The bases upon which the departmental examination processes were conducted (e.g., open, promotional, open and promotional)

To identify the specific departmental examination processes to be included in the cost comparison exercise, project staff reviewed a listing of departmental

examination processes conducted for the Office Technician (Typing) job classification from late 2000 through Spring 2002. Based upon this review and consideration of the above-noted criteria, three departmental examination processes for the Office Technician (Typing) classification were selected for inclusion in the cost comparison. The three examination processes utilized for cost comparison purposes had been conducted by the California Highway Patrol, the Department of Motor Vehicles, and the Department of the Youth Authority.

Data collection efforts focused on actual and estimated staff/resources costs, as well as timeline requirements, including the following:

- Staff time required to plan, develop, announce, schedule, and administer the examination process
- Staff time required to screen candidate applications
- Staff time required to prepare and distribute candidate notices
- Costs associated with the scanning of candidate answer sheets and the scoring of candidate test performance
- Facility costs for examination administration
- Supply and printing costs for the preparation of examination booklets, examination bulletins, and other testing materials

The cost comparison was conducted in May and June 2002.

### **State Personnel Board Hearing**

A hearing before the State Personnel Board was held on July 9, 2002, in conjunction with the State Personnel Board's regularly scheduled July Board meeting. The hearing served two purposes in the completion of this project. First, the hearing was used to solicit comments regarding the preliminary findings and recommendations of this project which were issued in June 2002 in preparation for the July 2002 Board hearing. Secondly, the hearing functioned as the final data collection element for this project. Public comment focusing on issues relevant to the evaluation of the State's decentralized testing program, including, but not limited to, the following was solicited:

- Problems or limitations with the decentralized testing program
- Advantages or positive attributes of the program
- Recommendations to improve the program

The 90-minute hearing was attended by approximately 100 individuals. Testimony and oral comments were provided by 14 individuals representing a variety of interests and groups, including unions, departments, individual employee/candidate interests, as well as the assessment profession.

In addition to the public hearing, public comment in the form of written testimony was also solicited. Several stakeholders provided written comments regarding the overall status of the State's decentralized testing program, as well as providing comments directly related to the preliminary findings and recommendations of the project.

A listing of the groups providing oral testimony during the hearing, as well as those groups providing written testimony in conjunction with the hearing, is provided in *Appendix H*. A copy of the hearing notice, without the Executive Summary attachment, is provided in *Appendix I*.

## **Results and Findings of Data Collection Efforts**

The following section of this report discusses in detail the results and findings of the data collection efforts documented herein.

### **Literature Review**

A review of the professional literature was conducted by project staff in early 2002. The purpose of the literature review was to identify research findings and published studies addressing decentralized testing practices within the human resources field. The review yielded the identification of several reports published by the Federal government documenting various facets of the Federal government's decentralized testing program.

Following are brief summaries of the relevant reports identified, as well as a brief overview of the findings and recommendations presented in each report:

#### ***The Role of Delegated Examining Units: Hiring New Employees in a Decentralized Civil Service***

**A report to the President and the Congress of the United States  
by the U.S. Merit Systems Protection Board, issued August 1999**

In August 1999, the United States Merit Systems Protection Board, an agency of the Federal government, issued a report detailing how Federal agencies, numbering nearly 700, were implementing the Federal government's decentralized testing program. The report was prepared as an initial evaluation of the effectiveness of the decentralized testing program, which was implemented in 1996. The report focused on the identification of the processes established by agencies to conduct their decentralized testing activities, the selection techniques used by agencies, favorable attributes of decentralized testing cited by agencies, and any concerns or deficiencies with the decentralized testing program. The main focus of the report was on initial appointments to the Federal civil service rather than other types of appointments involving existing Federal employees (e.g., promotions, transfers, reassignments).



The Federal government implemented decentralized testing to address the need for a faster, more flexible selection system that would facilitate the identification of best-qualified candidates to fill vacancies. Upon implementation of the decentralized testing program, the Office of Personnel Management (OPM), which had been the Federal government's centralized testing authority, was required to delegate examining authority to individual Federal agencies. OPM was then authorized to provide staffing assistance and conduct selection-related activities on behalf of other Federal agencies only on a reimbursable basis.

In reviewing the status of the Federal government's decentralized testing program, the report outlined the following main findings:

- Most decentralized testing processes are comprised of either a written examination or an unassembled examination comprised of an evaluation of one's experience, education, and training.
- Decentralized testing processes are faster and more effective in identifying high-quality candidates than were centralized testing processes.
- There is concern over the effectiveness of the evaluations of experience, education, and training in identifying qualified and best-qualified candidates.
- The "rule of three" certification rule often precludes the hiring authority from selecting the best-qualified candidate. Under the "rule of three," the hiring authority must select from the top three scoring candidates when filling a vacancy. The testing processes frequently fail to give the best qualified candidates the top scores in the testing process; thus, the hiring authorities are forced to select from lesser-qualified or not best-qualified candidates.
- Hiring supervisors and managers lack sufficient understanding of sound human resources principles and practices to make sound, merit-based selection decisions.

To improve the effectiveness of the Federal government's decentralized testing program, the report issued the following main recommendations:

- OPM should be allocated funds from Congress to develop valid written examinations which could be made available on a no-cost basis to agencies, such that agencies would be provided with better selection tools with which to assess candidate qualifications.
- Agencies need to commit adequate funds, staff, time, and training to the decentralized testing program to ensure that examination staff is adequately prepared to carry out their testing responsibilities and to ensure that candidates are afforded the opportunity to participate in fair, competitive testing processes.
- The application of the “rule of three” certification rule should be modified to allow agencies to determine on a case-by-case basis when it is appropriate to limit selection to only the top three scoring candidates and when it is appropriate to provide for the consideration of a greater number of candidates.
- Agencies should provide training to hiring supervisors and managers on human resources principles and practices, as specifically related to making sound, merit-based hiring decisions.

As of the writing of this report, these recommendations have yet to be fully implemented.

***Assessing Federal Job Seekers in a Delegated Examining Environment***  
**A report to the President and the Congress of the United States**  
**by the U.S. Merit Systems Protection Board, issued December 2001**

Similar to its report in August 1999, entitled *The Role of Delegated Examining Units: Hiring New Employees in a Decentralized Civil Service* and previously discussed herein, the United States Merit Systems Protection Board issued a report in December 2001, detailing how selection in the federal government's decentralized environment is conducted. The report outlined the primary selection means utilized by the nearly 700 federal agencies conducting decentralized testing. The report focused on hiring candidates into the Federal service. Promotions within the service were not addressed. Additionally, the report addressed preservation of the Federal merit system under provisions of the decentralized testing program and provided recommendations aimed at preventing breaches to the merit philosophy.

Over the past 20 years, the Federal government has moved toward a predominantly decentralized testing program. The impetus of the Federal government's shift to decentralization significantly mirrors the State's rationale for moving toward decentralization. As discussed previously, decentralization was fully implemented within the Federal service in 1996.

Under decentralization, OPM is charged with providing centralized testing authority for the Federal service. OPM has established standards and regulations with which agencies must comply when conducting their decentralized testing processes. In addition, OPM trains agency examination unit staff and periodically reviews and audits the decentralized testing processes. Actual selection work performed by OPM on behalf of other Federal agencies is done so on a reimbursable basis. Agencies, however, are under no obligation to seek services or assistance from OPM. Further, some agencies have reported that they lack sufficient resources to contract with OPM for services.

In reviewing the status of the Federal government's decentralized testing program, the report outlined the following main findings:

- The expertise of staff working in agencies with decentralized testing authority varies widely even though OPM provides training and oversight to all agencies engaged in decentralized testing.
- The value and importance of employment testing varies significantly across the agencies engaged in decentralized testing.
- Some agencies view employment testing as a cost rather than an investment, resulting in varying degrees of the quality of decentralized testing processes.
- The quality and soundness of decentralized testing processes are affected by each agency's funding and the amount of resources available to develop and use valid selection procedures.
- Decentralized testing affords agencies the opportunity to utilize testing processes that will best meet their specific selection needs in terms of timing, scheduling, and labor market response.
- A number of factors affect the quality and effectiveness of decentralized testing processes, including agency culture,

organizational structure, budget, staff expertise, and the number and types of job vacancies filled.

- A weakness in the rationale for moving toward decentralization was the expectation that individual agency testing efforts would maintain the quality and soundness of the Federal government's selection system, when in reality the quality of testing has suffered in some agencies as a result of decentralization.

In addressing the identified deficiencies of the Federal government's decentralized testing program, the report issued the following main recommendations:

- OPM should determine the degree of disparity between the requirement that all agencies use valid selection procedures and the extent to which some agencies fail to do so.
- OPM should develop a plan to eliminate any such disparities in the quality and soundness of decentralized testing processes.
- Agencies should view candidate assessment as a strategic business investment.
- Agencies should budget for the development and use of valid selection procedures.
- Agencies should provide for the adequate training of their examination unit staff.
- Agencies should hold managers, supervisors, and human resources staff responsible for sound selection.
- Agencies should balance the need for sound selection with the ability to hire the best-qualified candidates.

As of the writing of this report, these recommendations have yet to be fully implemented.

***The Federal Merit Promotion Program: Process vs. Outcome***  
**A report to the President and the Congress of the United States**  
**by the U.S. Merit Systems Protection Board, issued December 2001**

In December 2001, the United States Merit Systems Protection Board issued a report reviewing the Federal government's merit-based promotional process to ensure efficiency and effectiveness. The report outlined the main provisions of the Federal merit promotion program, under which nearly 100,000 promotions are made annually in the Federal service. The main focal points of the report were to determine whether promotions in the Federal service are based on merit, whether the merit promotion process provides for efficiency in making promotions, and the perception of Federal employees, supervisors, managers, and unions relative to the effectiveness of the merit promotion program. Additionally, the report addressed the amount of time and resources necessary to administer the merit promotion program.

Federal law requires that promotions in the Federal service "be determined solely on the basis of relative ability, knowledge, and skills, after fair and open competition which assures that all receive equal opportunity." The merit promotion program establishes the provisions and parameters under which promotions in the Federal service may be made. Individual agencies are required to establish and utilize a systematic means of affecting promotions on the basis of merit. Agencies are afforded latitude in determining the specific selection instruments to use in assessing candidate qualifications for promotional opportunities, provided that those selection techniques are valid and provide for merit-based selection.

In reviewing the status of the Federal government's merit promotion program, the report outlined the following main findings:

- Federal employees, supervisors, and union representatives all indicated reservations about the effectiveness of the merit promotion program in terms of allowing for promotion of the best-qualified individuals.
- There is a discrepancy regarding the perceived fairness of promotional decisions between supervisors, employees, and union representatives. For the most part, supervisors believe that promotional decisions are based on merit. To the contrary, employees and union representatives believe that promotions are

based on supervisor loyalty and “connections to other important people” in the Federal service.

- There is a belief among employees that pre-selection of candidates for promotion occurs.
- Supervisors have indicated that the identification of the best-qualified candidate can be more expediently and efficiently determined based on their personal knowledge of candidate qualifications than can occur through the use of the merit promotion process and formal selection instruments.
- The merit promotion process takes too long to fill vacancies.

Based on the review of the merit promotion process, the report issued the following main recommendations:

- Agencies should attempt to develop valid new selection procedures with which to assess candidate qualifications.
- Supervisors should be trained on how to develop and conduct structured interviews for use in the hiring process.
- Agencies should evaluate their individual merit promotion program provisions to ensure that process components are efficient.
- Agencies and OPM should explore the use of alternative non-competitive, merit-based selection procedures, such as promotions in place based on performance and the use of performance-based ascension through levels in classification series.

As of the writing of this report, these recommendations have yet to be fully implemented.

## **Survey of State Government Testing Practices**

A review of the testing practices of other state governments which utilize a merit-based selection model was conducted by project staff in early 2002. A total of 14 states were identified for survey purposes. Project staff was successful in reaching representatives from 12 of the 14 states.

In conducting the surveys, staff followed a structured methodology. After providing a brief explanation as to why the survey was being conducted, project staff used a standard set of questions to collect the following information:

- The extent to which each state utilizes decentralized testing
- For those states utilizing decentralized testing, the extent to which the decentralized testing programs are monitored for quality and effectiveness, and with whom the responsibility for the monitoring resides
- For states not utilizing decentralized testing programs, the extent to which the use of and/or the implementation of decentralized testing programs is being considered.

Following are the results of the survey.

Of the 12 states surveyed, seven (representing 58 percent of the sample) indicated that some form of decentralized testing is available to their respective state departments/agencies. All of the seven states which allow for some form of decentralized testing are merit-based, civil service systems, except for Georgia which is a merit-based, non-civil service system. Details of the main elements of these states' decentralized testing programs are presented as follows:

- Four states (Georgia, Maryland, Minnesota, and Wisconsin) allow decentralized testing for department-specific classifications; testing for multidepartmental and statewide classifications is conducted on a centralized basis by the central testing authority.
- One state (Michigan) indicated that departments have decentralized testing authority for all state classifications except for five of six "master" classifications for which testing is still currently conducted by the centralized testing authority. Under provisions of Michigan's decentralized testing program, departments are not allowed to develop formal testing procedures (e.g., written examinations, examination interviews, performance tests). Any formal testing procedures utilized on a decentralized basis must be developed by the Department of Civil Service, Michigan's centralized testing authority. Departments that choose not to utilize testing procedures developed by the Department of Civil Service are limited to using resume review processes only. Additionally, Michigan's Department of Civil Service is in the process of developing a

library of testing instruments which departments can opt to use in their decentralized testing processes.

- Two states (New York and Ohio) indicated minimal use of decentralized testing authority by state departments/agencies. New York indicated that while all departments could engage in decentralized testing, less than one percent of all state testing is conducted on a decentralized basis. The reason cited for the minimal use of decentralized testing is the fact that centralized testing services, which appear to meet departmental testing needs, are offered at no cost to the departments. Thus, it is believed that departments have little need or incentive to pursue the costly option of decentralized testing over the no-cost alternative of centralized testing. Ohio indicated that there is an effort underway within the state to recentralize all testing; therefore, departments are encouraged to not conduct decentralized testing processes. Within the past year, decentralized testing processes were conducted for only a couple of classifications within the state.

In terms of monitoring or auditing individual decentralized testing processes, three of the seven states with decentralized testing programs indicated that they currently utilize or will be implementing soon some type of audit/review process. An overview of each state's audit/review process follows:

- Maryland utilizes a process by which every position filled through a decentralized testing process is audited. The initial audit process consists of the appointing department documenting the process by which the position was filled, including whether the position was filled through the use of an existing eligible list or whether a new selection process was administered. If hiring process deficiencies are identified through a review of data submitted by the appointing department, a formal audit will be conducted by the centralized testing authority.
- Ohio's monitoring process requires departments to complete an audit checklist for every decentralized testing process conducted. The audit checklist is then reviewed by staff of the state's centralized testing authority. In the event that deficiencies are identified in any decentralized testing process, the department having conducted the process is notified in writing by the central testing authority of the deficiencies and admonished to refrain from committing such infractions in future testing processes. There is no follow-up or further action taken, however, if departments continue to conduct deficient testing processes.



- Georgia is in the process of implementing an on-line, self-rating audit tool. Using the self-rating tool, departments will self-certify the manner in which they have completed/conducted various aspects of their testing processes. The audit tool will be used by the departments to gauge how well they are conducting their testing processes and to identify areas of improvement for completion of future testing activities. There are no plans to have any central testing authority review or evaluate the completed audit tools.

None of the seven states with decentralized testing programs has conducted any type of program-wide review to determine the quality or effectiveness of the program. One state (Michigan) indicated that planning is underway to conduct such a systemic review/evaluation; however, no progress on those plans has been made to date.

In addition to the data collected in response to the standard questions, two respondents provided the following comments regarding their experience with decentralized testing programs:

- A major concern with decentralized testing programs is the duplication of effort when testing for servicewide or multidepartmental classifications, as well as when testing for general entry-level classifications.
- Unofficial review of decentralized testing indicates that it is a failure, with the greatest manipulation of the system occurring at the eligible list certification and hiring stages.
- One of rationales for decentralized testing is that individual departmental staff can be more responsive to the needs of the departments, which could be true at the expense of maintaining sound selection.
- By and large, decentralized testing is effective. However, it is absolutely clear that those departments with the less capable staff gravitate to the worst testing devices/techniques.

## **Review of Quality Assurance Audit Findings**

A review of audit findings documented by the SPB's Quality Assurance unit was conducted by project staff in December 2001. Project staff also interviewed the program manager to discuss and clarify audit findings, as well as to collect additional

information regarding the quality and effectiveness of the decentralized testing program.

Project staff reviewed data and findings from audit processes conducted for four departments. The departments audited were the Board of Prison Terms (BPT), the California Science Center (CSC), the Department of General Services (DGS), and the California Department of Veterans' Affairs (CDVA). It should be noted that the audits were specifically requested as a result of various concerns and issues with each of the departments' testing and/or personnel practices. The BPT audit was conducted in 2001 at the request of the Legislature in response to concerns and issues brought forward regarding BPT's personnel practices and specifically the examination process for the Deputy Commissioner, BPT classification. The BPT audit focused on examination activities from July 1, 1999, through June 30, 2001. The CSC audit was conducted in 2000 at the request of the State and Consumer Services Agency in response to findings and problems with CSC personnel practices as identified in an audit conducted by the Bureau of State Audits. The audit focused on CSC testing activities during calendar year 1999. The DGS audit was conducted as a result of a Board resolution calling for a review of DGS' testing practices between July 1997, and July 1999, to ensure compliance with applicable laws and regulations. The audit was conducted in 1999 and final results and findings were reported in 2001. The CDVA audit was conducted in 2001 at the request of California Senator John Burton as a result of issues and concerns regarding CDVA's personnel practices that were brought forward at the May 2000 Senate Confirmation Hearings of the former Secretary, CDVA. The CDVA audit focused on testing activities during 1999 and 2000.

These four audit reports provided data and findings for a total of 85 examination processes. The 85 examination process represented approximately one percent of the civil service examination processes conducted throughout the State during the total audit period.

Through review of the audit reports and findings, project staff identified a number of common deficiencies in the 85 examination processes reviewed/audited. The deficiencies cover a range of activities within the decentralized testing processes; thus, for ease in reporting, the deficiencies are categorized as follows:

#### Job-Relatedness of Examination Processes

The majority of examination processes reviewed lacked evidence of job-relatedness. There was little or no evidence or documentation linking performance in the examination process to levels of performance and/or qualifications required on the job. Without such evidence of job-relatedness, there is no guarantee that the

subject matter assessed or the qualifications which candidates were required to demonstrate were in any way linked to the requirements of the job classification for which the examination process was being conducted.

### Development of Examination Interview Processes

In many instances where interview processes were included as a component in the examination process, the interviews were not soundly developed. The development deficiencies included the content of the questions being asked, the number of questions being asked, and the manner in which the interviews were scored. In terms of the content of the questions, it was noted in the audits that frequently the interview questions were not appropriate in the examination setting and, in some instances, would have been better utilized as part of a hiring interview. The questions were not measuring candidate possession of specific job-related knowledge, skills, and abilities (KSAs). Typically, the deficient questions were alluding to candidate interest in the job classification and motivation to perform on the job.

In reviewing the number of questions comprising the interviews, it was noted in the audit findings that many of the interview processes lacked a sufficient number of questions to adequately assess candidate qualifications. In some instances, the interviews were comprised of as few as three questions. Further, in the majority of instances, the interviews failed to include at least seven questions or rating dimensions, which by professional standards is the minimum content requirement to ensure that the interview process provides an accurate, reliable measure of candidate qualifications.

### Scoring Criteria

The audit findings noted numerous deficiencies in the scoring criteria utilized for many of the examination interview processes. In some instances, no scoring criteria with which to rate candidate performance in the interview had been developed. Thus, there was no documented basis to indicate how or why candidates received the scores they were given. In addition, when scoring criteria were developed for the examination interviews, frequently the scoring criteria were overly subjective and relied on the interview panel's subjective interpretation of the criteria in order to rate candidate performance. The audit findings indicated that in these instances where candidate scores were based on overly subjective scoring criteria or no criteria at all, the examination interview processes were inadequate in assessing candidate qualifications. The audit findings further noted that such inadequacy of assessment resulted in examination processes that failed to provide for a competitive assessment of candidate qualifications.

### Administration and Scoring

Errors in the scoring of the interviews were detected in a number of the examination interview processes. With many of the interview processes, there were scoring errors resulting from the inaccurate application of established scoring criteria in rating candidate performance in the interviews. In some cases, the final scores assigned to candidates differed from the scores that had been recorded by the interview panel in their notes of the candidate's performance in the interview. For a number of interviews, candidate scores reflected scoring from panel members who had not actually been present in the interview and had not rated the candidates in actuality. Finally, in various processes, large scoring differentials existed among the panel members, wherein a single candidate was rated very high by one or more of the panel members and very low by one or the other panel members. Such differential scoring discrepancies, which diminish the accuracy with which the interview assesses candidate qualifications, are often a result of the use of poor scoring criteria which overly rely on panel member subjectivity in evaluating candidate performance or a result of the misapplication of established scoring criteria.

### Interpretation of Minimum Qualifications (MQs)

Additional deficiencies in the testing processes evaluated through the audit process included errors or discrepancies in the manner in which minimum qualification (MQ) requirements were applied and/or interpreted. In situations where MQ requirements were misapplied or misinterpreted, qualified candidates may have been excluded from the examination processes, while candidates failing to meet the MQ requirements were allowed to participate in such processes. Additionally, in some instances, candidates who were allowed to participate in examination processes clearly did not meet the MQ requirements. Finally, it was noted in the audit findings that for some examination processes reviewed, the MQ requirements were applied in an inconsistent manner. That is, the application screening of candidates with comparable experience, education, and/or background did not produce consistent screening results. In some of the examination processes audited, candidate applications could not be located for review. In other processes audited, candidate applications lacked pertinent information regarding the candidates' experience, education, and background. To the extent that the applications were not available for review and/or the applications lacked pertinent information, it could not be determined whether candidates allowed to participate in examination processes and those disqualified from participation actually met the MQ requirements. Such errors in interpreting and/or applying MQs were noted as being detrimental to the integrity of the testing process, having the potential to arbitrarily impact candidates seeking employment with the State, and possibly resulting in illegal appointments.

In addition to the review of the Quality Assurance audit findings, project staff also interviewed the program manager to collect additional data related to the quality and effectiveness of the decentralized testing program. Several issues were discussed in this interview and are documented as follows:

- Departments want flexibility to design and administer examination processes, as well as guidance and oversight related to testing policies and procedures from the SPB.
- Few departments are conducting job analysis, and as a result, the majority of examination processes conducted under the decentralized testing program would not meet the legal definition of being job-related.
- It is a disservice to the public and candidates when candidates are required to compete in multiple examination processes for the same job classification.
- Candidates are often unclear as to why they are required to compete in multiple departmental examination processes, rather than a single servicewide examination process, when they view the State as a single employer entity.
- The overlap in testing efforts when multiple departments test individually on open bases for servicewide classifications is inefficient.
- Departments often utilize different examination plans and/or testing instruments when testing for the same classification. (For example, one department may use a written examination to test for the Office Technician classification, and another department may use an interview process.)
- Departmental examination planning efforts and testing instrument selection often are not based on job-related criteria.
- Interpretation of the MQ requirements of a classification often differs from department to department, resulting in inappropriate, inaccurate, and inconsistent interpretation.
- There is a perception among employees and candidates that decentralized testing processes allow for manipulation.

## **Review of Examination Appeals Data**

A review of examination appeals records for fiscal years 1998/1999, 1999/2000, and 2000/2001 was conducted by project staff in December 2001. Project staff also interviewed the program manager to discuss and clarify the appeals documentation, as well as to collect additional information regarding recent appeals activities related to decentralized testing processes.

Under current law, several specific facets of the examination process may be appealed. Candidates may file examination appeals on the following bases:

- Discrimination in the examination process on the basis of protected group status (i.e., on the basis of race, gender, color, religion, national origin, physical or mental disability, age, political affiliation, ancestry, marital status, sexual orientation, or political or religious opinion)
- Fraud in the examination process
- Erroneous interpretation or application of minimum qualification requirements
- A significant irregularity in the examination process whereby the examination plan, as outlined on the examination bulletin, is not adhered to or the examination method is not applied fairly to all candidates

Review of examination appeals records and accompanying data for fiscal years 1998/1999, 1999/2000, and 2000/2001 resulted in the following findings.

### Fiscal Year 1998/1999

A total of 1,836 examination appeals were filed against 45 departments. As noted in Table 1, 1,286 (70%) of the appeals were filed against four departments. Those departments were the Department of Corrections, the Department of Transportation, the Employment Development Department, and the State Personnel Board. The remaining 550 appeals (30%) were filed against the other 41 departments. Forty-four (2.4%) of the appeals were granted.

Of the 299 appeals filed against the Department of Corrections, 246 (82%) were filed in three examination processes: 112 (37%) were filed in the Correctional Officer examination process, 40 (13%) were filed in the Parole Agent I, Adult examination process, and 94 (31%) were filed in the Special Agent, Corrections examination

**Table 1**  
***Appeals Data for Fiscal Year 1998/1999***

<b>Department</b>	<b>Examination Appeals Filed</b>	<b>% of Total Filed</b>	<b>Number (%) Granted</b>
Department of Corrections	299	16%	35 (12%)
Department of Transportation	431	23%	2 (0.5%)
Employment Development Department	312	17%	0 (0%)
State Personnel Board	244	13%	1 (0.4%)
Remaining 41 departments	550	30%	6 (1%)
<b>Total Examination Appeals</b>	<b>1,836</b>	<b>100%</b>	<b>44 (2.4%)</b>

process. Of the 35 appeals granted, six were granted in the Correctional Officer examination process. None were granted in the Parole Agent I, Adult, nor Special Agent, Corrections examination processes.

Of the 431 appeals filed against the Department of Transportation, 211 (49%) were filed in the Senior Transportation Engineer, Caltrans examination process. Neither of the two appeals granted were in the Senior Transportation Engineer, Caltrans examination process.

Of the 312 appeals filed against the Employment Development Department, 217 (70%) were filed in the Employment Program Representative examination process, and 58 (19%) were filed in the Employment Program Manager I examination process. None of the 312 appeals was granted.

Of the 244 appeals filed against the State Personnel Board, 88 (36%) were filed in the Office Assistant examination process, and 81 (33%) were filed in the Staff Services Analyst examination process. The one appeal granted was in the Staff Services Analyst examination process.

#### Fiscal Year 1999/2000

A total of 1,144 examination appeals were filed against 40 departments. As noted in Table 2, 631 (55%) of the appeals were filed against three departments. Those departments were the Department of Corrections, the Department of Transportation,



**Table 2**  
***Appeals Data for Fiscal Year 1999/2000***

<b>Department</b>	<b>Examination Appeals Filed</b>	<b>% of Total Filed</b>	<b>Number (%) Granted</b>
Department of Corrections	206	18%	2 (1%)
Department of Transportation	236	21%	1 (0.4%)
State Personnel Board	189	17%	0 (0%)
Remaining 41 departments	513	45%	4 (1%)
<b>Total Examination Appeals</b>	<b>1,144</b>	<b>100%</b>	<b>7 (0.6%)</b>

and the State Personnel Board. The remaining 513 appeals (45%) were filed against the other 37 departments. Seven (0.6%) of the appeals were granted.

Of the 206 appeals filed against the Department of Corrections, 116 (56%) were filed in two examination processes: 85 (41%) were filed in the Correctional Officer examination process, and 31 (15%) were filed in the Parole Agent II, Adult examination process. Of the two appeals granted, one was granted in the Correctional Officer examination process.

Of the 236 appeals filed against the Department of Transportation, 135 (57%) were filed in three examination processes: 84 (36%) were filed in the Caltrans Equipment Operator examination process, 33 (14%) were filed in the Senior Transportation Engineer, Caltrans examination process, and 18 (8%) were filed in the Transportation Engineering Technician examination process. The one appeal granted was not in any of these three examination processes.

Of the 189 appeals filed against the State Personnel Board, 79 (42%) were filed against the Staff Services Analyst examination process. None of the 189 appeals was granted.

#### Fiscal Year 2000/2001

A total of 1,089 examination appeals were filed against 59 departments. As noted in Table 3, 524 (48%) of the appeals were filed against three departments. Those departments were again the Department of Corrections, the Department of Transportation, and the State Personnel Board. The remaining 565 appeals (52%)



**Table 3**  
***Appeals Data for Fiscal Year 2000/2001***

<b>Department</b>	<b>Examination Appeals Filed</b>	<b>% of Total Filed</b>	<b>Number (%) Granted</b>
Department of Corrections	255	23%	0 (0%)
Department of Transportation	130	12%	0 (0%)
State Personnel Board	139	13%	2 (1%)
Remaining 41 departments	565	52%	6 (1%)
<b>Total Examination Appeals</b>	<b>1,089</b>	<b>100%</b>	<b>8 (0.7%)</b>

were filed against the other 56 departments. Eight (0.7%) of the appeals were granted. At the time of the file review, 74 (7%) appeals remained open under staff review.

Of the 255 appeals filed against the Department of Corrections, 124 (49%) were filed against two examination processes: 71 (28%) were filed against the Correctional Officer examination process, and 53 (21%) were filed against the Correctional Administrator examination process. None of these appeals was granted; however, 16 of the appeals filed against the Correctional Officer examination process and 50 of the appeals filed against the Correctional Administrator examination process remained open for staff review, for a total of 66 (26%) pending appeals.

Of the 139 appeals filed against the State Personnel Board, 53 (38%) were filed against the Staff Services Analyst examination process. Of the two appeals granted, one was granted in the Staff Services Analyst examination process. In addition, eight (6%) appeals filed against the State Personnel Board remained open under staff review.

#### Composite Data for Fiscal Years 1998/1999 through 2000/2001

A total of 4,069 examination appeals were filed against a total of 68 departments. As noted in Table 4, 2,441 (60%) of the appeals were filed against four departments. Those departments were the Department of Corrections, the Department of Transportation, the Employment Development Department, and the State Personnel Board. The remaining 1,628 (40%) appeals were filed against the other 64

**Table 4**  
***Composite Appeals Data for Fiscal Years 1998/1999 through 2000/2001***

<b>Department</b>	<b>Examination Appeals Filed</b>	<b>% of Total Filed</b>	<b>Number (%) Granted</b>
Department of Corrections	760	19%	37 (5%)
Department of Transportation	797	20%	3 (0.4%)
Employment Development Department	312	8%	0 (0%)
State Personnel Board	572	14%	3 (0.5%)
Remaining 41 departments	1,628	40%	16 (1%)
<b>Total Examination Appeals</b>	<b>4,069</b>	<b>100%</b>	<b>59 (1.4%)</b>

departments. Fifty-nine (1.4%) of the appeals were granted. At the time of the file review, 74 (2%) appeals remained open.

Additional findings as a result of the appeals data review include the following:

- The number of appeals filed in fiscal year 2000/2001 decreased by 41 percent compared to the number filed in fiscal year 1998/1999.
- The percentage of appeals granted in fiscal year 2000/2001, as of the time of the file review, decreased by 1.7 percent compared to the number of appeals granted in fiscal year 1998/1999.
- The Departments of Corrections and Transportation, as well as the State Personnel Board, account for a significant amount of appeals activity annually. These departments accounted for 53 percent of the appeals filed and 73 percent of the appeals granted during fiscal years 1998/1999 through 2000/2001.
- Of the 59 appeals granted during fiscal years 1998/1999 through 2000/2001, 37 (63%) were granted for examination processes conducted by the Department of Corrections, three (5%) were granted for examination processed conducted by the Department of Transportation, and three (5%) were granted for examination processes conducted by the State Personnel Board.

## Departmental Questionnaire

A comprehensive questionnaire was disseminated to the 89 State departments currently engaged in decentralized testing. The questionnaire elicited data regarding each department's testing activities for the fiscal years of 1998/1999, 1999/2000, and 2000/2001. A total of 77 (87%) completed questionnaires were returned. Questionnaire respondents recorded their responses directly in the questionnaire booklets. Project staff analyzed the questionnaire data using SPSS software. Table 5 illustrates the response rates by department size.

**Table 5**  
***Decentralized Testing Questionnaire Response Rates***

<b>Departments utilizing Decentralized Testing</b>	<b>Questionnaires Distributed</b>	<b>Questionnaires Returned</b>	<b>Response Rate</b>
Small Departments	60	50	83%
Medium Departments	12	11	92%
Large Departments	17	16	94%
<b>All Departments</b>	<b>89</b>	<b>77</b>	<b>87%</b>

The questionnaire was divided into several sections. The first section of the questionnaire elicited demographic data regarding each department's examination staff. The second section elicited data regarding each department's use of SPB products and services, and the third section elicited data regarding each department's decentralized examination activities for the fiscal years of 1998/1999, 1999/2000, and 2000/2001.

Analysis of the demographic section of the questionnaire produced the following results:

- Departments allocate an average of approximately seven positions to examination activities, not including time and resources spent for subject matter experts. Large departments allocate an average of 21.5 positions; medium departments allocate an average of six positions; and, small departments allocate an average of 2.45 positions.

- Approximately half of the departmental staff allocated to examination activities functions at the analyst level.
- The average ratio of examination supervisors/managers to staff, including analyst, technician, and clerical support staff, across departments is approximately 1 to 3.5. Large departments have an average ratio of approximately 1 to 4; medium departments have an average ratio of 1 to 6; and, small departments have an average ratio of 1 to 3.
- Approximately 60 percent of the State's examination staff, including managers, supervisors, analysts, technicians, and clerical support, has received some form of training/education in the area of Industrial/Organizational Psychology (I/O). Departments indicated that approximately 65 percent of the I/O training/education has been attained through training courses offered through the SPB's Technical Training Program. Approximately 46 percent of examination staff in large departments has received formal I/O training/education; approximately 79 percent of examination staff in medium departments has received formal I/O training/education; and, approximately 61 percent of examination staff in small departments has received formal I/O training/education.
- Approximately 15 percent of the State's examination staff has completed the SPB's Selection Analyst Training Program and been awarded certificates of completion. Approximately 14 percent of examination staff in large departments has completed the program; approximately 6 percent of examination staff in medium departments has completed the program; and, approximately 15 percent of examination staff in small departments has completed the program.
- Approximately 4 percent of the State's examination staff is active members in professional organizations, such as the Personnel Testing Council, the International Personnel Management Association, and the International Personnel Management Association Assessment Council. Approximately 3 percent of examination staff in large and medium departments is active in professional organizations, while approximately 7 percent of examination staff in small departments is active in professional organizations.

The second section of the questionnaire elicited data regarding each department's use of SPB products and services. Based on analysis of the data, all departments, regardless of size, cited the same five products/services as being of greatest value to their respective examination unit staffs. Those five products are as follows:

- Consultation with SPB staff from various units, including Examination Services, Policy Consultation, Quality Assurance, Test Validation and Construction, Appeals, and Information Services.
- The *Selection Manual*
- Courses offered through the Technical Training Program
- The On-Line Certification/Examination training provided through the Examination Services Unit
- SPB's website

The third section of the questionnaire elicited data regarding each department's decentralized examination activities for the fiscal years of 1998/1999, 1999/2000, and 2000/2001. Analysis of the data collected produced the following results.

### **Fiscal Year 2000/2001**

A total of 3,968 examination processes were conducted in fiscal year 2000/2001, and approximately 209,329 candidates competed in those examination processes. Of those examination processes, 49 percent were conducted on an open basis, 41 percent were conducted on a promotional basis, 9 percent were conducted on an open, non-promotional basis, and the remaining 0.5 percent were conducted on an open and promotional basis. Approximately 61 percent of the examination processes were conducted by large departments; approximately 17 percent were conducted by medium departments; and, approximately 23 percent were conducted by small departments.

### **Instrument Usage**

A variety of selection instruments were used in the examination processes conducted. While a number of instruments were utilized, the predominant method of candidate assessment was the examination interview, comprising 59 percent of the selection instruments administered. In addition, the rates of usage of some instruments differed by department size. Small departments utilized examination interviews less than did large and medium departments; however, small department utilized ratings of education, training, and experience (E&Es) far more than did medium and large departments. Table 6 illustrates the various selection instruments utilized and the percentage of use of each instrument.

**Table 6**  
***Selection Instrument Usage, Fiscal Year 2000/2001***

Instrument	Percentage of Use			
	All Departments	Large Departments	Medium Departments	Small Departments
Examination interviews	59%	60%	69%	49%
Ratings of education, training, and experience (E&Es)	18%	15%	16%	29%
Written examinations	10%	14%	7%	3%
Supplemental applications	7%	6%	7%	10%
Performance tests	2%	3%	0.7%	2%
Promotional Readiness Examinations (PREs)	2%	2%	0%	5%
Statements of qualifications	1%	0.3%	0.1%	2%

### Examination Planning

A number of the examination processes conducted in fiscal year 2000/2001 utilized multiple testing components to assess candidate qualifications (e.g., a written examination and an examination interview, a performance test and an examination interview). Of the 3,968 examination processes conducted, 415 (10%) were comprised of multiple testing components. Of the 2,419 examination processes conducted by large departments, 286 (12%) were comprised of multiple testing components; of the 656 examination processes conducted by medium departments, 40 (6%) were comprised of multiple testing components; and, of the 893 examination processes conducted by small departments, 89 (10%) were comprised of multiple testing components.

### Examination Development Activities

Approximately 736 of the examination processes conducted in fiscal year 2000/2001 were developed for first-time administration. These processes were developed on the bases of various job analytic approaches, including a review of the classification

specification, SME input, and formal job analysis. A combination of classification specification review and SME input was the approach used to develop the largest percentage of examination processes across all departments. However, the rates of usage of the various job analytic approaches differed by department size. Large departments utilized formal job analysis techniques in far greater proportion than did medium and small departments. Table 7 presents the percentage of examination processes developed based upon the four job analytic approaches.

**Table 7**  
***Job Analytic Approaches, Fiscal Year 2000/2001***

Basis	Percentage of Use			
	All Departments (n=736)	Large Departments (n=462)	Medium Departments (n=114)	Small Departments (n=160)
Review of classification specification and SME input	76%	69%	93%	83%
Job analysis	22%	31%	2%	11%
Review of classification specification only	2%	0%	5%	6%
SME input only	0.3%	0.2%	0%	0.6%

Of the examination processes newly developed for administration in fiscal year 2000/2001, some were developed by departmental examination staff and others were developed on a consultative basis by SPB staff and outside consultants. Table 8 presents the percentage of examination processes developed by the various examination development sources.

### **Fiscal Year 1999/2000**

A total of 3,296 examination processes were conducted in fiscal year 1999/2000, and approximately 209,240 candidates competed in those examination processes. Of those examination processes, 48 percent were conducted on an open basis, 41 percent were conducted on a promotional basis, 9 percent were conducted on an open, non-promotional basis, and the remaining one percent were conducted on an open and promotional basis. Approximately 66 percent of the examination processes were conducted by large departments; approximately 16 percent were

**Table 8**  
**Sources of Examination Development, Fiscal Year 2000/2001**

Source of Examination Development	Percentage of Use			
	All Departments (n=736)	Large Departments (n=462)	Medium Departments (n=114)	Small Departments (n=160)
Departmental Examination Staff	90%	88%	97%	95%
Departmental Subject Matter Experts (with no assistance from departmental examination staff)	3%	3%	1%	4%
SPB's Test Validation and Construction Unit	3%	5%	2%	0.6%
SPB's Examination Services Unit	3%	4%	0%	0%
Outside Consultants	0.5%	0.6%	0%	0.6%

conducted by medium departments; and, approximately 18 percent were conducted by small departments.

### Instrument Usage

A variety of selection instruments were used in the examination processes conducted. While a number of instruments were utilized, the prevalent method of candidate assessment was the examination interview, comprising 60 percent of the selection instruments administered. In addition, the rates of usage of some instruments differed by department size. Small departments utilized E&Es and supplemental applications to a greater extent than large and medium departments, while large and medium departments utilized examination interviews and written examinations more so than small departments. Table 9 illustrates the various selection instruments utilized and the percentage of use of each instrument.

### Examination Planning

A number of examination processes conducted in fiscal year 1999/2000 utilized multiple testing components to assess candidate qualifications. Of the 3,296



**Table 9**  
***Selection Instrument Usage, Fiscal Year 1999/2000***

Instrument Type	Percentage of Use			
	All Departments	Large Departments	Medium Departments	Small Departments
Examination interviews	60%	62%	64%	47%
Ratings of education, training, and experience (E&Es)	17%	14%	17%	29%
Written examinations	10%	12%	9%	3%
Supplemental applications	9%	8%	7%	14%
Performance tests	2%	3%	1%	2%
Promotional Readiness Examinations (PREs)	2%	2%	0.2%	5%
Statements of qualifications	0.3%	0.2%	0.5%	0.5%

examination processes conducted, 355 (11%) were comprised of multiple testing components. Of the 2,164 examination processes conducted by large departments, 263 (12%) were comprised of multiple testing components; of the 541 examination processes conducted by medium departments, 59 (11%) were comprised of multiple testing components; of the 591 examination processes conducted by small departments, 33 (6%) were comprised of multiple testing components.

### Examination Development Activities

Approximately 626 of the examination processes conducted in fiscal year 1999/2000 were developed for first-time administration. These processes were developed on the bases of various job analytic approaches, including a review of the classification specification, SME input, and formal job analysis. A combination of classification specification review and SME input was the approach used to develop the largest percentage of examination processes across all departments. However, the rates of usage of the various job analytic approaches differed by department size. Large departments utilized formal job analysis techniques in far greater proportion than did medium and small departments. Table 10 presents the percentage of examination processes developed based upon the four job analytic approaches.

**Table 10**  
***Job Analytic Approaches, Fiscal Year 1999/2000***

Basis	Percentage of Use			
	All Departments (n=626)	Large Departments (n=417)	Medium Departments (n=120)	Small Departments (n=89)
Review of classification specification and SME input	82%	76%	96%	92%
Job analysis	17%	24%	0%	8%
Review of classification specification only	0.8%	0.3%	3%	0%
SME input only	0.2%	0%	1%	0%

Of the examination processes newly developed for administration in fiscal year 1999/2000, some were developed by departmental examination staff and others were developed on a consultative basis by SPB staff and outside consultants. Table 11 presents the percentage of examination processes developed by the various examination development sources.

### **Fiscal Year 1998/1999**

In responding to the questionnaires, a number of departments indicated that their records did not include complete data and documentation for fiscal year 1998/1999. Therefore, in drawing conclusions based upon these data, it must be noted that while the actual numbers may not reflect the total amount of examination activity conducted, the trend of examination activities appears to be stable.

Based upon the data submitted, a total of 2,660 examination processes were conducted in fiscal year 1998/1999, and approximately 200,851 candidates competed in those examination processes. Of those examination processes, 41 percent were conducted on an open basis, 46 percent were conducted on a promotional basis, 11 percent were conducted on an open, non-promotional basis, and the remaining one percent were conducted on an open and promotional basis. Approximately 65 percent of the examination processes were conducted by large

**Table 11**  
**Sources of Examination Development, Fiscal Year 1999/2000**

Source of Examination Development	Percentage of Use			
	All Departments (n=626)	Large Departments (n=417)	Medium Departments (n=120)	Small Departments (n=89)
Departmental Examination Staff	88%	92%	94%	88%
Departmental Subject Matter Experts (with no assistance from departmental examination staff)	4%	3%	0%	12%
SPB's Test Validation and Construction Unit	0.3%	0.4%	0%	0%
SPB's Examination Services Unit	7%	3%	23%	0%
Outside Consultants	1%	2%	0%	0%

departments; approximately 16 percent were conducted by medium departments; and, approximately 20 percent were conducted by small departments.

### Instrument Usage

A variety of selection instruments were used in the examination processes conducted. While a number of instruments were utilized, the main method of candidate assessment was the examination interview, comprising 52 percent of the selection instruments administered. In addition, the rates of usage of some instruments differed by department size. Small departments utilized written examination to a much lesser extent than did large and medium departments, while utilizing E&Es to a greater degree than large and medium departments. Table 12 illustrates the various selection instruments utilized and the percentage of use of each instrument.

A number of examination processes conducted in fiscal year 1998/1999 utilized multiple testing components. Of the 2,660 examination processes reportedly conducted, 352 (13%) were comprised of multiple testing components. Of the 1,724 examination processes conducted by large departments, 244 (14%) were comprised

**Table 12**  
***Selection Instrument Usage, Fiscal Year 1998/1999***

Instrument Type	Percentage of Use			
	All Departments	Large Departments	Medium Departments	Small Departments
Examination interviews	52%	52%	61%	43%
Ratings of education, training, and experience (E&Es)	20%	19%	11%	35%
Written examinations	11%	13%	12%	3%
Supplemental applications	10%	10%	13%	11%
Performance tests	3%	3%	1%	2%
Promotional Readiness Examinations (PREs)	3%	3%	2%	5%
Statements of qualifications	0.2%	0.3%	0%	0%

of multiple testing components; of the 413 examination processes conducted by medium departments, 65 (16%) were comprised of multiple testing components; and, of the 523 examination processes conducted by small departments, 43 (8%) were comprised of multiple testing components.

#### Examination Development Activities

Approximately 510 of the examination processes reportedly conducted in fiscal year 1998/1999 were developed for first-time administration. These processes were developed on the bases of various job analytic approaches, including a review of the classification specification, SME input, and formal job analysis. A combination of classification specification review and SME input was the approach used to develop the largest percentage of examination processes across all departments. As the rates of usage of the job analytic approaches differed by department size, it should be noted that large departments utilized formal job analysis techniques to a much greater extent than did medium and small departments. Table 13 presents the percentage of examination processes developed based upon the four job analytic approaches.

**Table 13**  
***Job Analytic Approaches, Fiscal Year 1998/1999***

Basis	Percentage of Use			
	All Departments (n=510)	Large Departments (n=327)	Medium Departments (n=98)	Small Departments (n=85)
Review of classification specification and SME input	86%	80%	97%	94%
Job analysis	13%	19%	1%	6%
Review of classification specification only	0.6%	0.3%	2%	0%
SME input only	0.2%	0.3%	0%	0%

Of the examination processes newly developed for administration in fiscal year 1998/1999, some were developed by departmental examination staff and others were developed on a consultative basis by SPB staff and outside consultants. Table 14 presents the percentage of examination processes developed by the various examination development sources.

### **Trends during Fiscal Years 1998/1999 through 2000/2001**

Based upon the data collected in the questionnaire, the following findings and trends related to the decentralized testing program during fiscal years 1998/1999 through 2000/2001 are reflected:

#### Overall Examination Activity

- The number of examination processes conducted has increased approximately 20 percent each year.

The proportion of examination processes conducted by large departments, medium departments, and small departments has remained relatively unchanged with the largest changes occurring between fiscal year 1999/2000 and fiscal year

**Table 14**  
**Sources of Examination Development, Fiscal Year 1998/1999**

Source of Examination Development	Percentage of Use			
	All Departments (n=510)	Large Departments (n=327)	Medium Departments (n=98)	Small Departments (n=85)
Departmental Examination Staff	86%	89%	77%	82%
Departmental Subject Matter Experts (with no assistance from departmental examination staff)	6%	6%	0%	12%
SPB's Test Validation and Construction Unit	0.4%	0.6%	0%	0%
SPB's Examination Services Unit	7%	2%	21%	6%
Outside Consultants	1%	1%	1%	0%

2000/2001. In fiscal year 2000/2001, the proportion of examination processes conducted by large departments decreased by approximately 5 percent, and the proportion conducted by small departments increased by approximately 5 percent. The proportion of processes conducted by medium departments remained relatively unchanged during this time.

- The number of candidates competing in the State's examination processes has remained relatively stable with only a 4 percent increase in the number of candidates competing in 2000/2001 examination processes compared to the number reported to have competed in 1998/1999 processes.
- The percentage of examination processes which allow for competition on an open basis has remained relatively stable. An increase of 5 percent was noted in the percentage of processes conducted on an open basis between 1998/1999 and 1999/2000. The percentage of processes allowing for open competition increased a mere one percent between 1999/2000 and 2000/2001. Further, testing within large and medium departments provides for a far greater degree of testing on an open basis than do the testing efforts of small departments.

- The use of multiple components in examination processes has decreased slightly in past three fiscal years. Based on the data reported for examination activities in fiscal year 1998/1999, 13 percent of the examination processes conducted were comprised of multiple components. That figure decreased to 10 percent of the processes conducted in fiscal year 2000/2001. This decrease was most significantly affected by the activities of medium departments. In fiscal year 1998/1999, 16 percent of the processes conducted by medium departments were multiple-component processes, whereas in fiscal year 2000/2001, only 6 percent of the processes were multiple-component processes.

#### Selection Instrument Usage

- In reviewing the types of instruments used by departments, the rate of usage for all instrument types has remained fairly constant during fiscal years 1998/1999 through 2000/2001 for the total group of departments. Further, the rate of usage by department size has fluctuated little.
- Examination interviews have been the most prevalently used instrument, accounting for more than 50 percent of the instruments administered in each fiscal year.

#### Job-Relatedness of Examination Processes

- The use of job analysis in the development of examination processes increased by 9 percent between fiscal years 1998/1999 and 2000/2001, and by 5 percent between fiscal years 1999/2000 and 2000/2001.
- Departments indicated having conducted job analysis for 22 percent of the examination processes developed in fiscal year 2000/2001. Further, large departments indicated that job analysis was conducted for 31 percent of the examination processes developed in fiscal year 2000/2001.

#### Examination Development

- The development of examination processes has been predominately conducted by departmental examination staff. During fiscal years 1998/1999 and 1999/2000, little change was noted in the departmental use of the SPB's Test Validation and Construction (TV&C) unit and the Examination Services unit for examination development activities. However, between fiscal years 1999/2000 and 2000/2001, reliance on the TV&C unit for examination development services increased nearly ten-fold, while reliance on the Examination Services unit

decreased by more than 50%. The significant users of TV&C unit services are large departments, for which the TV&C unit developed approximately 5 percent of the new examination processes conducted by large departments during fiscal year 2000/2001.

- Departmental use of professional consultants outside of the State has remained minimal over the past three fiscal years, accounting for less than 1 percent of all newly developed examination processes during this time period.

### Correlational Analysis of the Data

Correlation analysis of the demographic data set and examination activity data from fiscal year 2000/2001 was conducted to determine if there were any relationships between departmental staffing and decentralized testing activities. Table 15 presents the correlational matrix. The major findings based upon the correlational analysis can be summarized as follows:

- The amount of I/O training/education possessed by the departmental examination staff correlates significantly with department size.
- The number of departmental examination staff who have completed all of the courses offered through the SPB's Selection Analyst Training Program (SATP) correlates significantly with department size and the size of the departmental examination staff.
- The use of multiple-component examination processes correlates significantly with the amount of I/O training/education possessed by departmental examination staff, departmental examination staff involvement in human resources professional organizations, and departmental examination staff completion of the SATP program.

### **Departmental Interviews**

Interviews with a representative sampling of 16 of the 89 State departments (18%) currently engaged in decentralized testing were conducted by project staff in March and April 2002. The purpose of the interviews was to collect data regarding specific departmental testing practices, as well as departmental perceptions of the effectiveness and quality of decentralized testing.



**Table 15**  
**Correlations among Staffing and Decentralized Testing Activities**

	n	Correlations								
		1	2	3	4	5	6	7	8	9
1. Department size	77									
2. Number of candidates	76	.56**								
3. Size of examination staff	77	.47**	.90**							
4. I/O training/education	76	.46**	.84**	.97**						
5. Completion of SATP	71	.40**	.72**	.82**	.84**					
6. Professional organizations	71	.18	.41**	.19	.12	.14				
7. Job analysis	26	.37	.37	.20	.15	.11	.07			
8. Multi-component processes	74	.52**	.74**	.48**	.41**	.40**	.64**	.38		
9. Number of appeals filed	69	.42**	.77**	.87**	.82**	.77**	.09	.32	.40**	
10. Number of appeals granted	46	.26	.84**	.95**	.91**	.72**	.13	.31	.33*	.81**

Note: For department size, 0 = small department, 1 = medium department, 2 = large department.

\*\* - < .01. \* - < .05.

In conducting the interviews, staff followed a structured methodology which included a brief overview of the project and the use of a standardized interview comprised of 10 questions. The interview addressed the job-relatedness of the decentralized testing processes conducted by the departments, the extent to which training or guidance is provided by the departments to hiring supervisors, departmental perceptions of the effectiveness of the decentralized testing program, as well as the role of the SPB in the decentralized testing program, and the level of departmental participation in the SPB's Selection Analyst Training Program and corresponding application of the concepts and principles presented in the training classes. The interview questions are provided in *Appendix D*.

The responses to the interview questions were analyzed across the 16 departments for their thematic content. Following are the results of the interviews presented by content area:

## **Job-relatedness of decentralized testing processes**

Questions one through three identified departmental practices related to establishing the job-relatedness of examination processes. The questions focused on the extent to which job analyses are conducted, the extent to which SMEs are utilized in examination development activities, and the process by which examination plans are developed.

With regard to conducting job analyses, nine departments (56%) conduct job analysis to some degree. Four departments (24%), all of which are large departments, indicated that formal job analyses are conducted for all of the examination processes developed. Five departments (31%), two of which are small departments, two of which are medium departments, and one of which is a large department, indicated that job analyses are conducted for some, but not all, of the examination processes developed. Finally, seven departments (44%), four of which are small departments, and three of which are medium departments, indicated that no job analyses are conducted for their examination processes.

The 12 departments (75%) not conducting job analyses for all of their examination processes provided various reasons for omitting this step. The primary reason cited by eight of the 12 departments (67%) is a lack of resources in terms of both staff time and expertise. The other four departments (33%) indicated that since either they test primarily on a promotional basis or the classifications for which they test have not changed recently, it is not necessary to conduct job analyses.

With regard to the use of SMEs in the design and development of examination processes, all 16 departments utilize SMEs in this regard. In addition, all of the departments indicated that adequate numbers of SMEs are provided to examination staff when designing and developing examination processes. In terms of the level of SMEs utilized, twelve departments (75%) select SMEs who are incumbents in the job classification for which the testing process is being conducted and/or first- or second-level supervisors of incumbents in the job classification. The other four departments (25%) select SMEs who are at least two levels above the job classification for which the testing process is being conducted.

In response to how examination plans are developed, 15 of the 16 departments (94%) indicated that examination staff work in conjunction with SMEs and program management to determine the testing instruments to comprise the examination process and the accompanying scoring model. One small department indicated that the examination plan from the previous administration of the examination process is reviewed by the Personnel Officer who determines the new examination plan.

### **Supervisor training/guidance regarding hiring interviews/hiring decisions**

With respect to training provided to hiring supervisors in conducting hiring interviews and making hiring decisions, 11 departments (69%) indicated that support is provided to hiring supervisors. This support follows a variety of formats. Four (36%) of the 11 departments provide formal training to hiring supervisors, addressing the hiring process, how to conduct hiring interviews, and how to make hiring decisions. Five (45%) of the 11 departments disseminate to hiring supervisors handbooks and other materials which address conducting hiring interviews and making hiring decisions. Two (18%) of the 11 departments indicated that human resources staff provides assistance to hiring supervisors in developing hiring interview questions and conducting the interviews.

The remaining five departments interviewed (31%), four of which are small departments, and one of which is a large department, indicated that currently no such training or formal support is provided to hiring supervisors. However, the one large department indicated that a manual on this topic is currently being developed for hiring supervisors but is not yet available.

### **The effectiveness of the decentralized testing program**

A series of questions addressing the aspects of the decentralized testing program that are working, the aspects of the program that are currently not working, and elements of the program that can be improved was asked. The majority of departments provided very similar comments in response to these questions. The main aspects of the decentralized testing program cited as currently working include the following:

- 13 departments (81%) indicated that decentralized testing processes can be conducted more timely than centralized processes could be.
- 12 departments (75%) indicated that departments have greater flexibility with decentralized testing to the extent that they can utilize more diverse, state-of-the-art examination processes than they might otherwise have available to them under a centralized testing system.
- 10 departments (63%) indicated that departmental selection needs can be responded to more efficiently through decentralized testing processes. Examination processes can be prioritized based on departmental needs rather than the systemic needs of a centralized testing system. Hiring can

be completed faster using decentralized testing processes than using centralized processes.

- Six departments (38%) indicated that departmental examination staff are more familiar with the classifications for which examination processes are conducted since they are employed in the department. This familiarity has resulted in increased satisfaction in the examination process for program staff and a greater degree of cooperation from them.

The main elements of the decentralized testing program cited as currently not working include the following:

- Ten departments (63%) cited a number of deficiencies with the SPB's on-line examination scoring system which directly impact departments' ability to process decentralized examinations. Deficiencies cited include the system being cumbersome to use, incapable of processing complex examination scoring models, and antiquated in terms of requiring overnight processing time frames.
- Three departments (38%) indicated that inconsistent information and/or inconsistent interpretations of policies, rules, and regulations provided by SPB staff adversely impact the effectiveness of the decentralized testing program.

In response to the questions of how the centralized testing program could be improved and what the SPB could do to improve the program, the departments cited the following program improvements:

- Thirteen departments (81%) indicated that the SPB's *Selection Manual* should be updated to provide better and more current guidance and direction to departmental examination staff.
- Ten departments (63%) responded that the SPB's on-line examination scoring system should be updated to provide more timely processing of data and a greater capability to process complex examination scoring models.
- Nine departments (56%) stated that the SPB should provide better customer service in terms of staff disseminating consistent information and interpretation of policies, rules, and regulations, as well as providing more timely response to telephone calls and e-mails. Departments also

indicated that the SPB should be allocated additional staff to provide for better customer services.

### **The role of the SPB in the decentralized testing program.**

Questions related to the perception of the SPB's role in the decentralized testing program generated the following responses from 12 (75%) responding departments:

- The SPB should function as a control agency, establishing testing-related expectations and standards for departments and holding departments accountable.
- The SPB should function as a support agency, providing technical expertise, expert resources, and guidance to departments engaged in testing activities, as well as leading the State in cutting-edge, innovative testing processes.
- The SPB should provide training to departmental examination staff related to conducting sound, effective, and innovative testing processes.

### **Departmental participation in the SPB's Selection Analyst Training Program**

The interview questions regarding the extent to which examination staff attend training courses offered through the SPB's SATP program provided the following data:

- Departmental examination staff in 13 departments (81%) have attended training courses offered in the SATP program. Further, all of the examination staff in 11 of those 13 departments (85%) have attended one or more of the courses.
- The primary reasons cited by departments as precluding a greater degree of participation in the SATP program training courses are a lack of or limited financial resources and insufficient staff time/availability to attend.

The interview questions regarding the extent to which examination staff apply the concepts and principles presented in the training classes provided the following data:

- Twelve of 13 departments (92%) that have sent staff to one or more of the training classes offered through the Selection Analyst Training Program

indicated that staff are applying concepts and principles learned. The 12 departments include five small departments, three medium departments, and four large departments. One medium department indicated that staff have attended classes too recently to have yet utilized the concepts and principles learned.

- The primary concepts and principles cited by the 12 departments as being applied to departmental examination activities are presented in Table 16. The two most widely applied training class concepts are the use of structured interviews, including structured scoring criteria, and job analysis.

**Table 16**  
***Frequency of Application of Training Concepts and Principles***

Concept/Principle	Frequency			
	All Departments (n=12)	Large Departments (n=4)	Medium Departments (n=3)	Small Departments (n=5)
Use of structured interviews	9	3	3	3
Use of job analysis	7	4	1	2
Examination planning	3	--	1	2
Restricting review of applications by interview panels	2	1	--	1
Pass point methodology	2	1	1	--
Overall process soundness	2	--	2	--

## Union Interviews

Interviews with representatives of eight of the 12 unions representing the State's 21 bargaining units (67%) were conducted by project staff in March and April 2002. The purpose of the interviews was to collect data and comments regarding the effectiveness of the decentralized testing program.

In conducting the interviews, staff followed a structured methodology which included a brief overview of the project and the use of a standardized interview comprised of

three questions. The interview addressed perceptions of the effectiveness of the decentralized testing program, as well as identified improvements that the SPB could make to the program. The interview questions are provided in *Appendix F*.

The responses to the interview questions were analyzed across the eight unions for their thematic content. Following are the results of the interviews presented by content area:

### **The effectiveness of the decentralized testing program**

A series of questions addressing the aspects of the decentralized testing program that are working, the aspects of the program that are currently not working, and elements of the program that should be discontinued was asked. The majority of union representatives provided very similar comments in response to these questions. The main aspects of the decentralized testing program cited as currently working include the following:

- Three unions (38%) indicated that decentralized testing processes can be conducted more timely and more frequently than centralized processes could be.
- Two unions (25%) indicated that departmental selection needs in terms of localized hiring can be responded to more efficiently through decentralized testing processes. Recruitment efforts and examination processes can be focused on particular geographic areas, resulting in a greater likelihood of filling vacancies in specific locations with interested, qualified candidates.

It should be noted that two unions (25%) indicated that no aspects of the decentralized testing program are currently working.

The main elements of the decentralized testing program cited as currently not working include the following:

- Six unions (75%) indicated a multitude of deficiencies in examination interview processes. The deficiencies included the inclusion of non-job-related questions; the use of questions geared for a specific candidate's qualifications rather than the requirements of the classification for which the interview was being conducted; lack of standardized scoring criteria; subjective assessment of candidate qualifications by interview panels; interview panels comprised of individuals lacking the technical expertise to evaluate candidate qualifications; the use of different questions for

different candidates; panel member bias contaminating interview processes; and, the use of panel members who are related and/or familiar to some of the candidates.

- Six unions (75%) cited a number of deficiencies in departmental communication with candidates and the notification of various examination process components provided to candidates. Deficiencies cited include failure to notify candidates when they are reachable on eligible lists or when vacancies in classifications for which they are eligible to be considered occur; failure to notify candidates of their appeal rights; failure to notify candidates of their right to remove panel members from examination interview panels; failure to adequately publicize examination bulletins and examination opportunities; and, a lack of feedback to candidates on their performance in examination processes.
- Five unions (63%) indicated that widespread favoritism in State service results in hires and appointments being made not on the basis of merit but, rather, on the basis on “who” candidates know.
- Four unions (50%) indicated that certification rules (e.g., rule of three names, rule of three ranks) and methods of scoring examinations (e.g., 9-limited scores, 6-limited scores, full range scoring) are not applied consistently in like or similar examination processes which can be detrimental to candidates and preclude consideration of qualified candidates in the hiring process.
- Three unions (38%) indicated the presence of pre-selection, occurring prior to examination processes being conducted and candidate qualifications actually being known.
- Three unions (38%) indicated that some departmental examination staff lack testing expertise which directly affects the quality and effectiveness of the examination processes conducted by these staff and impacts the degree of merit present in the State's selection system.
- Two unions (25%) indicated a lack of timeliness with decentralized testing processes, stating that decentralized testing processes take as long to administer as centralized processes used to and that the frequency with which some examination processes are conducted is less than when the processes were administered on a centralized basis.



- Two unions (25%) indicated a lack of department accountability when poor, non-merit-based examination processes are conducted and when non-merit-based hiring decisions are made.

The one element of the decentralized testing program cited by two unions (25%) as needing to be discontinued is the practice of using interview panel members who are familiar with the candidates. In addition, three unions (38%) indicated that while some improvements may be necessary, there are no specific elements of the program that should be discontinued.

### **Improvements to the decentralized testing program**

In response to the question of what improvements the SPB could make to the decentralized testing program, the unions cited the following program improvements:

- Five unions (63%) indicated that the SPB should provide more oversight and review of decentralized testing processes to ensure the soundness, merit, and job-relatedness of the processes, including holding departments accountable for compliance with SPB rules and regulations. One union further indicated that the SPB should mandate the validation of all testing processes.
- Two unions (25%) indicated that the SPB should mandate training for departmental examination staff before staff is authorized to develop and administer examination processes.
- Two unions (25%) stated that the SPB should provide for more timely processing and resolution of examination appeals.

### **Employee Advocacy Group Interviews**

Interviews with representatives from two employee advocacy groups were conducted by project staff in April 2002. The purpose of the interviews was to collect data and comments regarding the effectiveness of the decentralized testing program.

In conducting the interviews, staff followed a structured methodology which included a brief overview of the project and the use of the same standardized interview as was used for the union interviews. The interview questions are provided in *Appendix F*.

The responses to the interview questions were analyzed between the two advocacy groups and compared to the responses provided by the eight unions for their thematic content. Following are the results of the interviews presented by content area:

### **The effectiveness of the decentralized testing program**

In response to a series of questions which addressed the aspects of the decentralized testing program that are working, the aspects of the program that are currently not working, and elements of the program that should be discontinued, the advocacy groups provided comments similar to those provided by the unions.

Analogous to the input provided by the unions, the advocacy groups cited timely examination administration and geographically localized testing as the main aspects of the decentralized testing program that are currently working.

The main elements of the decentralized testing program cited as currently not working include the following, which correspond to those elements identified by the unions:

- Deficiencies in the examination interview process, including the use of interview panel members who are familiar with the candidates
- Deficiencies in departmental communication with candidates, including failure to adequately publicize examination bulletins and examination opportunities and a failure to publicize eligible lists
- Favoritism and pre-selection that impact the types of testing instruments utilized in examination processes, such that instruments that will favor specific candidates' qualifications are used; and, the weighting of testing instruments that comprise the examination process, such that those instruments that most favor specific candidates' qualifications carry the greatest, if not all, weight in determining final scores in the process
- The inconsistent application of certification rules and scoring methods
- The lack of testing expertise of some departmental examination staff

Also cited by the advocacy groups as a problematic element in the program is a general reluctance of candidates to utilize the appeals process or challenge an examination process for fear of reprisal.

The one element of the decentralized testing program cited as needing to be discontinued is the practice of using departmental staff to chair examination interview panels.

### **Improvements to the decentralized testing program**

In response to the question of what improvements the SPB could make to the decentralized testing program, the advocacy groups cited the following program improvements, which parallel those identified by the unions:

- The SPB should provide more oversight and review of decentralized testing processes, including holding departments accountable for compliance with SPB rules and regulations.
- The SPB should mandate training for departmental examination staff before staff is authorized to develop and administer examination processes.

An additional improvement cited called for the SPB to promote professionalism in the State's selection system by setting standards and job-related requirements for positions utilized to conduct examination activities.

### **Review of Decentralized Testing Processes**

A review of 86 decentralized testing processes was conducted by project staff in March and April 2002. The review process utilized standardized evaluation criteria to collect objective data pertaining to the quality, soundness, and fairness of a sampling of decentralized testing processes. The 86 examination processes reviewed included 29 processes conducted by five large departments, 29 conducted by five medium departments, and 28 conducted by six small departments. The examination processes represented a sampling of open and promotional testing processes. Table 17 illustrates the testing bases for the examination processes reviewed.

In addition to categorizing the examination processes on the basis of competition (e.g., open, promotional), the sampling of examination processes were also categorized on the basis of their scoring models to ensure that examination processes incorporating a variety of scoring models were evaluated. The sample of 86 examination processes included 67 processes (78%) comprised of a single component and 19 processes (22%) comprised of multiple components, some of

**Table 17**  
**Competition Bases of Decentralized Testing Processes Reviewed (n = 86)**

Basis of Examination	All Departments (n = 86)		Large Departments (n = 29)		Medium Departments (n = 29)		Small Departments (n = 28)	
	#	% of total	#	% of total	#	% of total	#	% of total
Open	29	34%	4	14%	13	45%	12	43%
Promotional	43	50%	18	62%	12	41%	13	46%
Open, Non-Promotional	12	14%	6	21%	3	10%	3	11%
Open and Promotional	2	2%	1	3%	1	3%	0	--

which were weighted, and some of which carried no weight in determining final candidate scores. Table 18 illustrates the various scoring models represented in the examination processes reviewed. A listing of the specific selection instruments comprising the scoring models of the examination processes reviewed is provided in *Appendix J*.

**Table 18**  
**Scoring Models of Decentralized Testing Processes Reviewed (n = 86)**

Scoring Model	All Departments (n = 86)		Large Departments (n = 29)		Medium Departments (n = 29)		Small Departments (n = 28)	
	#	% of total	#	% of total	#	% of total	#	% of total
Single component	67	78%	21	72%	23	79%	23	82%
Multiple weighted components	13	15%	5	17%	3	10%	5	18%
Multiple weighted and non-weighted components	6	7%	3	10%	3	10%	0	--

Included in the 86 examination processes reviewed were a variety of individual selection instruments. A total of 19 written examinations were reviewed, 64

examination interviews, 15 performance tests, two supplemental applications, and nine E&E processes. Table 19 illustrates the types and numbers of selection instruments reviewed by department size.

**Table 19**  
***Types of Instruments Reviewed (n = 109)***

<b>Types of Instruments</b>	<b>All Departments</b> (n = 109)		<b>Large Departments</b> (n = 39)		<b>Medium Departments</b> (n = 36)		<b>Small Departments</b> (n = 34)	
	#	% of total	#	% of total	#	% of total	#	% of total
Written examinations	19	17%	12	31%	4	11%	3	9%
Examination interviews	64	59%	20	51%	23	64%	21	62%
Performance tests	15	14%	4	10%	6	17%	5	15%
Supplemental applications	2	2%	1	3%	1	3%	0	--
E&Es	9	8%	2	5%	2	6%	5	15%

The standardized evaluation criteria used to evaluate the 86 decentralized testing processes addressed a multitude of variables relative to the design, development, and administration of each of the processes evaluated. The evaluation criteria utilized a scoring process consisting of a Likert-type scale with values ranging from 0 to 3. The criteria and corresponding scale were used to measure several elements indicative of minimal acceptable assessment practices, including job-relatedness, degree of structure, appropriateness, and completeness. The lower bound of the scale (0) reflected the presence of less-than-acceptable criteria. This rating was reserved for those instances where no attempt or an inadequate attempt was made to meet minimal standards. A rating of 1 indicated less-than-acceptable criteria but recognized an attempt to meet minimal standards in terms of job-relatedness, structure, completeness, and appropriateness. A rating of 2 reflected acceptable criteria and was indicative of practices and processes that embodied minimal acceptable assessment standards. The scale's upper bound (3) indicated greater-than-acceptable criteria and was utilized for those variables where evidence of practices and processes in excess of minimal standards could be noted. Table 20 presents the evaluation categories, as well as the individual rating dimensions in each category and the scoring range established for each dimension. A copy of the complete evaluation criteria is provided in *Appendix G*.

**Table 20**  
***Evaluation Criteria Categories***

<b>Evaluation Categories</b>	<b>Rating Dimensions</b>	<b>Scoring Range</b>
<b>Job-Relatedness</b>	Job analysis	0 – 3
	Job analytic data	0 – 3
	SME participation	0 – 2
<b>Examination Planning</b>	Examination base	0 – 2
	Instrument selection	0 – 3
	Scoring model	0 – 3
	SME participation	0 – 2
<b>Written Examination Development</b>	Development methodology	0 – 3
	SME participation in development	0 – 2
	Instrument content	0 – 2
	Item construction	0 – 3
	Preliminary pass point setting activities	0 – 3
	Pass point setting methodology	0 – 3
<b>Interview Development</b>	Development methodology	0 – 2
	SME participation in development	0 – 2
	Instrument content	0 – 3
	Question construction	0 – 3
	Scoring Criteria	0 – 3
	Pass point setting methodology	0 – 2
<b>Performance Test Development</b>	Development methodology	0 – 3
	SME participation in development	0 – 2
	Instrument content	0 – 3
	Instrument construction	0 – 3
	Scoring criteria	0 – 3
	Preliminary pass point setting activities	0 – 3
	Pass point setting methodology	0 – 3
<b>Supplemental Application Development</b>	Development methodology	0 – 2
	SME participation in development	0 – 2
	Instrument content	0 – 3
	Question/item construction	0 – 3
	Scoring criteria	0 – 3
	Pass point setting methodology	0 – 3
<b>Examination Administration</b>	Candidate instructions, as applicable	0 – 3
	Rater training/preparation, as applicable	0 – 3
	Rater qualifications	0 – 2
	Application of scoring criteria	0 – 2
	Availability of examination statistics	0 – 2
	Use of examination statistics	0 – 2

For ease in presenting the data, the results of the review process are presented categorically based upon the elements of the scoring criteria.

### Job-Relatedness of Examination Processes

The majority of examination processes lacked evidence of job-relatedness, and few were based on a formal job analysis. Of the 86 examination processes reviewed, 19 (22%) were based on job analyses that met the job analytic requirements set forth by the *Uniform Guidelines*. As such, these 19 job analyses constituted legally defensible, professionally acceptable job analyses. Further, these job analyses could represent the legal basis to substantiate the content validity of the examination processes subsequently developed. Four examination processes (5%) were based on job analyses that failed to conform to the requirements set forth in the *Uniform Guidelines*. The remaining 63 examination processes (73%) lacked any evidence of job analysis. The 19 job analyses that met minimal acceptable standards were conducted by seven of the 16 departments. The seven departments included four large departments, which conducted 10 (53%) of the job analyses, one medium department, which conducted one (5%) job analysis, and two small departments, which conducted eight (42%) of the job analyses.

The next rating dimension evaluated the presence of job analytic data, defined as the collection of job data that, while not in complete conformance with a recognized job analysis methodology, could serve to establish the job-relatedness of an examination process. The intent of the job analytic data criterion was to determine the extent to which jobs were analyzed in terms of identifying and documenting duties and corresponding KSA requirements. A distinction in the evaluation criteria was made between an attempt to analyze the job and simply the use of the classification specification as the basis for any task/KSA analysis. In those instances where the classification specification was the basis for the job analysis process, the resulting processes did not meet the minimal threshold for job-relatedness as defined in the evaluation criteria. In utilizing this rating dimension, project staff determined that those 19 processes which had met the standard of a formal job analysis would not be rated on this dimension. Thus, the remaining 67 processes were rated. The four processes noted above as having been based on job analyses which did not meet the definition of a sufficient job analysis were deemed to meet the minimal threshold for appropriate collection of job analytic data under this rating dimension. These four job analytic data collection efforts were completed by four large departments. Sixty-one (91%) of the remaining processes were based on attempts to collect job analytic data; however, the data collected and methodology followed in these instances did not meet minimal documentation requirements to establish the job-relatedness of the subsequently developed testing process. The two remaining processes (3%) were based on no attempt to collect job analytic data and, therefore, lacked documentation and evidence of job-relatedness.

The final dimension in this category evaluated the extent to which SMEs were appropriately involved in the job analytic process. Overall, SMEs were used sufficiently in the collection and/or attempted collection of job analytic data. Seventy-one processes (83%) met the minimum standard for appropriate SME use. Five processes (6%) attempted to utilize SMEs adequately but failed to meet the minimal standard, and 10 processes (12%) completely failed to utilize SMEs in the job analytic process. Table 21 illustrates the evaluation results in the Job-Relatedness category.

**Table 21**  
***Evaluation Results – Job-Relatedness***

Rating Dimension	n	Mean	Median	Frequency Results	
				Rating	Frequency
Job analysis	86	.55	.00	3	5 (6%)
				2	14 (16%)
				1	4 (5%)
				0	63 (73%)
Job analytic data	67	1.03	1.00	3	0 (0%)
				2	4 (6%)
				1	61 (91%)
				0	2 (3%)
SME participation	86	1.71	2.00	2	71 (83%)
				1	5 (6%)
				0	10 (12%)

Correlational analysis of the data was conducted to determine if there were any relationships between the job-relatedness of decentralized testing processes and departmental staffing activities. Table 22 presents the correlational matrix based on the data collected via this examination review process and general demographic data for the 16 departments that conducted the 86 examination processes reviewed. The departmental demographic data was collected via the *Decentralized Testing Questionnaire* described previously in this report. The major findings based upon correlational analysis can be summarized as follows:

- Job analysis correlates significantly with departmental examination staff size.
- Job analysis correlates significantly with the I/O training and expertise of departmental examination staff.



**Table 22**  
**Correlations among Staffing and the Job-Relatedness**  
**of Decentralized Testing Processes**

		Correlations					
	n	1	2	3	4	5	6
1. Department size	86						
2. Size of examination staff	80	.48**					
3. Training through SATP	86	.31**	.41**				
4. I/O training/education	80	.48**	.99**	.43**			
5. Professional organizations	80	-.41**	.11	.21	.08		
6. Job Analysis	86	.17	.35**	.35**	.33**	.43**	
7. Job analytic data	67	.27**	-.25	.05	-.28*	-.37**	.82**

*Note:* For department size, 1 = small department, 2 = medium department, 3 = large department. For training through SATP (SPB's Selection Analyst Training Program), 0 = no staff attendance in SATP classes, 1 = some or all staff have attended some or all classes, but no one is certified, 2 = some staff SATP-certified, 3 = all staff SATP-certified.

\*\* - < .01. \* - < .05.

- Job analysis correlates significantly with departmental examination staff participation in the SPB's SATP program.
- Job analysis correlates significantly with departmental examination staff involvement in human resources professional organizations.

### Examination Planning

A review of examination planning activities for the 86 examination processes evaluated indicated that for a large majority of the examination processes, departments had engaged in appropriate examination planning and design activities. Included in the examination planning activities reviewed were consideration of the use of appropriate examination bases upon which to test (e.g., open, promotional), use of selection instruments appropriate for the KSAs to be assessed, use of

appropriate scoring models based upon the job-related weighting of candidate scores in the examination process, and the use of SMEs in the planning process.

All of the processes reviewed utilized appropriate examination bases, such that the selection needs at hand could adequately be met as a result of the basis of competition. In evaluating the extent to which appropriate selection instruments were chosen to comprise the 86 processes, 77 processes (90%) utilized instruments which would adequately measure the KSAs identified for assessment. Nine processes (10%) were comprised of instruments that were not suited to measure the KSAs indicated for assessment. An additional evaluation factor was the appropriateness of selected scoring models. Of the 86 processes reviewed, 75 processes (87%) incorporated scoring models that were based upon a job-related rationale of quantifying candidate qualifications and, as developed, would provide for the appropriate weighting of candidate scores in the examination process. Eleven (13%) of the processes failed to incorporate the use of job-related scoring models. Based on available documentation related to the role of SMEs in the examination planning process, 74 processes (86%) appropriately incorporated SME input.

Overall, examination planning activities for the majority of processes reviewed were adequate. Table 23 illustrates the evaluation results in the Examination Planning category.

**Table 23**  
***Evaluation Results – Examination Planning***

Rating Dimension	n	Mean	Median	Frequency Results	
				Rating	Frequency
Examination base	86	2.00	2.00	2	86 (100%)
				1	0
				0	0
Instrument selection	86	2.18	2.00	3	26 (30%)
				2	51 (59%)
				1	7 (8%)
				0	2 (2%)
Scoring model	86	2.09	2.00	3	30 (35%)
				2	45 (52%)
				1	0
				0	11 (13%)
SME participation	86	1.80	2.00	2	74 (86%)
				1	6 (7%)
				0	6 (7%)

Correlational analysis of the data was conducted to determine if there were any relationships between the quality of examination planning activities and departmental staffing activities, as well as departmental testing practices. Table 24 presents the correlation matrix based on the data collected via this examination review process and general demographic data for the 16 departments that conducted the 86 examination processes reviewed. The major findings based upon correlational analysis can be summarized as follows:

- Instrument selection correlates significantly with departmental examination staff involvement in human resources professional organizations
- Scoring models correlate significantly with instrument selection.

**Table 24**  
***Correlations among Staffing and the Examination Planning Activities***  
***for Decentralized Testing Processes***

		Correlations					
	n	1	2	3	4	5	6
1. Department size	86						
2. Size of examination staff	80	.48**					
3. Training through SATP	86	.31**	.41**				
4. I/O training/education	80	.48**	.99**	.43**			
5. Professional organizations	80	-.41**	.11	.21	.08		
6. Instrument selection	86	.10	.17	.19	.17	.31**	
7. Scoring model	86	.05	.14	.20	.15	.28*	.57**

*Note:* For department size, 1 = small department, 2 = medium department, 3 = large department. For training through SATP (SPB's Selection Analyst Training Program), 0 = no staff attendance in SATP classes, 1 = some or all staff have attended some or all classes, but no one is certified, 2 = some staff SATP-certified, 3 = all staff SATP-certified.

\*\* - < .01. \* - < .05.

### Written Examination Development

Nineteen (22%) of the 86 examination processes reviewed incorporated the use of a written examination. The written examination component was either utilized as the sole selection instrument or one of multiple instruments comprising the process. A number of factors related to the development of the written examinations were

evaluated for the 19 written examinations, including the following:

- The methodology followed in developing the written examination
- The content of the instrument based upon the KSAs identified for assessment and the extent to which test segments of adequate length and item homogeneity were developed
- The construction of the individual items in the test, including wording, stem clarity, and distracter plausibility
- Pass point setting activities
- The use of SMEs in the development process

The strongest development facets of the 19 written examinations reviewed were in the areas of development methodology, pertaining to the job-relatedness of the content, and item construction. Seventeen examinations (89%) met or exceeded the minimum criteria established for development methodology which called for instrument development on the basis of SME participation and some form of job analytic data. Based on available documentation related to the role of SMEs in the examination development process, 16 examinations (84%) appropriately incorporated SME input. In terms of the quality of the individual items comprising the written examinations, 16 examinations (84%) were comprised of items that for the most part were well-constructed and in adherence to established professional item writing standards.

With respect to the overall content design of the examinations, eight of the examinations (42%) failed to meet basic content parameters. The predominant content deficiency in these eight examinations was a failure to incorporate adequate numbers of items in individual segments within the examination, per recognized testing principles. Most of these examinations were comprised of one or more segments with fewer than 20 items in the segment.

The most noticeable deficiency present in the 19 examinations reviewed was inadequate pass point setting activities. Two facets of pass point setting were evaluated. The first facet focused on preliminary pass point setting activities and the collection of job-related data upon which to base the pass points that would ultimately be set for these instruments. The evaluation criteria called for the use of a professionally recognized methodology to collect the job-related pass point data. While eight of the examinations (42%) reviewed did include the collection of job-related data upon which to base subsequent pass points, the remaining 11 examinations (58%) failed to even attempt the collection of such job-related data.

The eight examinations for which preliminary pass point setting data were collected had been conducted by four large departments. It should be noted that three of these eight written examinations had been developed and validated by the SPB's Test Validation and Construction (TV&C) unit, and TV&C staff were actively engaged in the collection of the preliminary pass point setting data.

The second facet of pass point setting which was evaluated was the actual pass point setting process itself and the degree to which the pass points set for these examinations could be substantiated as job-related. The evaluation criteria were structured to identify the use of a pass point methodology that would yield defensible results in the examination process. Use of an acceptable methodology would result in pass points that when applied could differentiate between those candidates possessing the required qualifications and those who did not. Factors considered in evaluating the presence or lack of an acceptable pass point methodology were the use of job-related data as a basis for the pass point, examination statistics, and the consideration of the examination's validity. Of the 19 examinations reviewed, only four (21%), all of which were conducted by large departments, had job-related, defensible pass points. It should again be noted that three of these four written examinations were developed and validated by the SPB's Test Validation and Construction (TV&C) unit under contract with three of the four large departments, and TV&C staff were actively engaged in the pass point setting process for these examinations. The other 15 examinations (79%) failed to utilize sound pass points. The pass points ultimately set were arbitrary and not based on job-related expectations of performance on the examinations. Of these 15 examinations which failed to utilize job-related pass points, eight (53%) were conducted by large departments, four (27%) by medium departments, and three (20%) by small departments.

Table 25 illustrates the evaluation results in the Written Examination Development category.

#### Written Examination Administration

In addition to evaluating the development of the 19 written examinations, various administration activities of those instruments were reviewed, including the following:

- The instructions provided to candidates via the test booklets
- The availability of examination statistics
- The use of examination statistics in reviewing item performance and setting pass points

**Table 25**  
***Evaluation Results – Written Examination Development***

Rating Dimension	n	Mean	Median	Frequency Results	
				Rating	Frequency
Development methodology	19	2.42	3.00	3	10 (53%)
				2	7 (37%)
				1	2 (11%)
				0	0
SME participation	19	1.79	2.00	2	16 (84%)
				1	2 (11%)
				0	1 (5%)
Instrument content	19	1.58	2.00	2	11 (58%)
				1	8 (42%)
				0	0
Item Construction	19	2.26	2.00	3	8 (42%)
				2	8 (42%)
				1	3 (16%)
				0	0
Preliminary pass point setting activities	19	1.11	.00	3	5 (26%)
				2	3 (16%)
				1	0
				0	11 (58%)
Pass point setting methodology	19	1.32	1.00	3	4 (21%)
				2	0
				1	13 (68%)
				0	2 (11%)

- The extent to which the administration of the examinations resulted in adverse impact.

Review of the test booklets and the instructions provided to candidates indicated that all 19 written examinations had provided clearly worded, standardized candidate instructions.

Deficiencies were noted in the availability and use of various examination statistics. Availability of a variety of reports was evaluated due to the importance of those reports in finalizing and setting pass points for written examinations. The specific

reports identified in the review criteria included item analyses, raw score tabulation reports, examination control documentation, and bottom line hiring reports. In reviewing the files and documentation for the 19 written examinations, the specified reports were available for 14 (74%) of the 19 processes; however, the reports were not available for the other five processes (26%).

In addition to reviewing the files of the 19 written examinations to determine the presence of the aforementioned reports and the relevant examination statistics, the files were also reviewed to determine whether the data had been appropriately utilized by examination staff in finalizing the written examinations, including identifying and deleting defective items and setting the pass points. It was noted that in fewer than half of the examination processes had the examination statistics and specified reports been utilized correctly. In eight of the processes (42%), all of which were conducted by large departments, the reports and statistics had been used appropriately, while in the other 11 processes (58%), the reports and statistics had not been used appropriately. These 11 processes were conducted by departments of varying size; four of the processes were conducted by large departments, four by medium departments, and three by small departments. For some of these 11 processes, the data was available, but departments did not utilize it. In some cases, had the examination statistics been interpreted and applied accordingly, poor-performing examination items could have been deleted from the examination.

Table 26 illustrates the evaluation results in the Written Examination Administration category.

A review of adverse impact data for the 19 written examinations revealed that eight (42%) of the 19 examinations resulted in adverse impact against one or more candidate groups. Of those eight occurrences of adverse impact, three occurrences (38%) could be defended on the basis of business necessity, as those three examinations had validation evidence documenting their job-relatedness. The examinations in the other five occurrences had not been validated nor developed on the basis of job analysis, and, therefore, conclusions about whether the content of the examinations was related to business necessity could not be determined.

### Examination Interview Development

Sixty-four (74%) of the 86 examination processes reviewed incorporated the use of an examination interview. The interview component was either utilized as the sole selection instrument or one of multiple instruments comprising the process.

**Table 26**  
***Evaluation Results – Written Examination Administration***

Rating Dimension	n	Mean	Median	Frequency Results	
				Rating	Frequency
Candidate instructions	19	2.79	3.00	3	0
				2	19 (100%)
				1	0
				0	0
Availability of examination statistics and corresponding reports	19	1.78	2.00	2	14 (74%)
				1	5 (26%)
				0	0
Use of examination statistics and corresponding reports	19	1.44	1.00	2	8 (42%)
				1	11 (58%)
				0	0

A number of factors related to the development of the 64 examination interviews were evaluated, including the following:

- The methodology followed in developing the interview
- The content of the instrument based upon the KSAs identified for assessment, including the types of questions used (e.g., situational, behavioral consistency, experience, job knowledge) and the number of questions/rating dimensions
- The construction of the individual questions
- The scoring criteria
- The pass point setting methodology
- The use of SMEs in the development of the instrument.

In reviewing the 64 examination interview processes, the strongest development facets of the interviews were in the areas of development methodology, pertaining to the job-relatedness of the content, and question construction. Fifty-five interviews (86%) met or exceeded the minimum criteria established for development methodology which called for instrument development on the basis of SME



participation and some form of job analytic data. Based on available documentation, 59 interviews (92%) appropriately incorporated SME input in the development process. An evaluation of the quality of the individual interview questions found that 61 interviews (95%) met or exceeded the minimum quality standard. The quality facet consisted of an assessment of the question types appropriate for the KSAs being measured, as well as the clarity with which the questions were worded questions.

A number of deficiencies were identified in the 64 examination interviews reviewed. The deficiencies included interview content, scoring criteria, or lack thereof, and the establishment of pass points. Several factors with respect to interview content were evaluated. Those factors included the length of the interview, the types of questions used, and the specific KSAs identified for assessment. Only 26 interviews (41%) met the content design requirements, while 38 interviews (59%) failed to meet those requirements. The predominant deficiency in these 38 interviews was a failure to utilize an adequate number of questions in the interview process. A minimum of seven questions or rating dimensions was considered the threshold to ensure assessment of an adequate sample of candidate qualifications. Most of the 38 interviews deficient in this area were comprised of fewer than seven questions or rating dimensions, with many of the interviews being comprised of as few as four questions/rating dimensions. Of the 26 interviews which adequately met content design requirements, 10 (38%) were conducted by large departments, nine (35%) by medium departments, and seven (27%) by small departments. Of the 38 interviews which failed to achieve content design requirements, 14 (37%) were conducted by large departments, 14 (37%) by medium departments, and 10 (26%) by small departments.

A second serious deficiency noted in half of the examination interviews evaluated was the scoring process and accompanying scoring criteria utilized for the interviews. To meet the minimum threshold of acceptability, a structured, standardized, job-related scoring process was needed to rate candidate responses. Thirty-two of the interviews (50%) reviewed met or exceeded this threshold; the other 32 interviews (50%) failed to do so. Many of these interviews that lacked adequate scoring criteria had little or no documented criteria for the interview panels to use in rating candidate performance and, instead, relied on the subjective judgement of the interview panels to evaluate candidate qualifications. Of the 32 interviews with adequate scoring criteria, 10 (31%) were conducted by large departments, 13 (41%) by medium departments, and nine (28%) by small departments. Of the 32 interviews which failed to utilize adequate scoring criteria, 10 (31%) were conducted by large departments, 10 (31%) by medium departments, and 12 (38%) by small departments.

The final area of deficiency noted in the examination interview processes reviewed was the degree to which the pass points in the interviews represented a job-related rationale with which to determine whether candidates adequately possessed the KSAs being assessed. The interviews were reviewed to determine if a pass point or threshold of passing had been established in the scoring criteria. That point of passing or threshold was then evaluated in terms of its job-relatedness and the degree to which it was adhered to by the interview panel in the scoring of candidate performance in the interview. Of the 64 interviews reviewed, 34 (53%) utilized adequate, documented, job-related pass points. Of those 34 processes, 11 (32%) were conducted by large departments, 12 (35%) by medium departments, and 11 (32%) by small departments. The other 30 interviews (47%) failed to utilize sound pass points. Nine (30%) of those 30 interviews were conducted by large departments, 11 (37%) by medium departments, and 10 (33%) by small departments.

Table 27 illustrates the evaluation results in the Interview Development category.

**Table 27**  
***Evaluation Results – Interview Development***

Rating Dimension	n	Mean	Median	Frequency Results	
				Rating	Frequency
Development methodology	64	1.86	2.00	2	55 (86%)
				1	9 (14%)
				0	0
SME participation	64	1.91	2.00	2	59 (92%)
				1	4 (6%)
				0	1 (2%)
Interview content	64	1.52	1.00	3	7 (11%)
				2	19 (30%)
				1	38 (59%)
				0	0
Question Construction	64	2.44	2.00	3	31 (48%)
				2	30 (47%)
				1	3 (5%)
				0	0
Scoring criteria	64	1.61	1.50	3	23 (36%)
				2	9 (14%)
				1	16 (25%)
				0	16 (25%)
Pass point setting methodology	64	1.25	2.00	2	34 (53%)
				1	12 (19%)
				0	18 (28%)

### Examination Interview Administration

In addition to evaluating the development of the 64 examination interview processes, various administration activities of those instruments were reviewed, including the following:

- The instructions provided to candidates if the interview processes included some type of work sample exercise/pre-interview exercise administered in conjunction with the interview
- The extent to which interview panel members were trained/prepared for their assessor roles
- The qualifications of the panel members
- The application of the scoring criteria by the panel members in rating candidate performance in the interviews
- The availability of examination statistics
- The extent to which the administration of the interviews resulted in adverse impact

Fifteen (23%) of the 64 examination interview processes included the use of a pre-interview exercise, comprised of a work sample, performance-based exercise and/or pre-exposed interview questions. Review of the instructions for the pre-interview exercises provided to candidates indicated that 14 (93%) of the pre-interview exercises had provided clearly worded, standardized instructions for the candidates. The one remaining pre-interview exercise (7%) failed to provide clearly worded instructions to the candidates; however, instructions were provided that attempted to meet minimal acceptable standards.

Based on available records, 63 (98%) of the 64 examination interview processes provided for sufficient training and preparation of the interview panels prior to administration of the interview processes. Documentation regarding rater training was not available for one of the interview processes reviewed; therefore, the presence and/or adequacy of such training could not be determined. The qualifications of the interview panel members was also a facet of the evaluation process. Documentation of panel member qualifications in 61 of the processes (95%) indicated that the panel members were appropriately qualified to evaluate candidate qualifications. For the other three interview processes (5%), documentation of rater qualifications was not sufficient to determine the appropriateness of the raters.

To determine the extent to which the scoring criteria or established scoring process had been applied by the interview panels, a review of panel members' notes, audio recordings of a sample of interviews, final scoring documentation, and actual scoring criteria was completed. This review focused on the application of the interview scoring criteria as opposed to the quality of the scoring criteria. A comparison was made among the various records and documentation to determine the extent to which the interviews had been scored in a consistent manner. Of the 62 examination interview processes reviewed, documentation for 50 (81%) of the processes indicated consistent scoring practices by the interview panels. Seventeen (34%) of these 50 interviews were conducted by large departments, 15 (30%) by medium departments, and 18 (36%) by small departments. There were scoring inconsistencies noted in the remaining 12 interview processes (19%). The scoring inconsistencies noted in this review process occurred exclusively in those interviews for which inadequate scoring criteria had been developed. Of these 12 interview processes, three (25%) were conducted by large departments, six (50%) by medium departments, and three (25%) by small departments.

Additional deficiencies noted in the examination interview processes were the availability of various examination statistics and corresponding reports. The specific reports identified in the review criteria included examination control documentation and bottom line hiring reports. In reviewing the files and documentation for the 64 interview processes, it was determined that the specified reports were available for 49 (77%) of the 64 processes. The reports were not available for the other 15 processes (23%). In many of these instances, the departments which had conducted the interview processes do not currently use the SPB's on-line examination scoring system, from which these reports are typically generated.

Table 28 illustrates the evaluation results in the Interview Administration category.

Adverse impact data were available for 54 of the 64 examination interview processes reviewed. Documentation/data were not available for 10 of the interview processes. Of the 54 examination interview processes for which documentation/data were available, six interview processes (11%) resulted in adverse impact against one or more candidate groups. Of those six occurrences of adverse impact, three occurrences (50%) could be defended on the basis of business necessity, as those three examinations had validation evidence documenting their job-relatedness. The interviews in the other three occurrences had not been validated nor developed on the basis of job analysis, and, therefore, conclusions about whether the content of the interviews was related to business necessity could not be determined.

**Table 28**  
**Evaluation Results – Interview Administration**

Rating Dimension	n	Mean	Median	Frequency Results	
				Rating	Frequency
Candidate instructions	15	2.67	3.00	3	11 (73%)
				2	3 (20%)
				1	1 (7%)
				0	0
Interview panel training/ preparation	63	2.52	3.00	3	33 (52%)
				2	30 (48%)
				1	0
				0	0
Interview panel qualifications	64	1.95	2.00	2	61 (95%)
				1	3 (5%)
				0	0
Application of scoring criteria	62	1.78	2.00	2	50 (81%)
				1	4 (6%)
				0	8 (13%)
Availability of examination statistics and corresponding reports	64	1.63	2.00	2	49 (77%)
				1	6 (9%)
				0	9 (14%)

Additional analysis of the data revealed the following:

- The more departments utilize the structured interview model (comprised of job-related questions and standardized, structured scoring criteria), the more sound the resulting examination interview processes.
- The efficiencies and deficiencies in the examination interview processes reviewed are fairly uniformly proportioned across departments of varying size.
- The more departmental examination staff participate in the SPB's SATP program, the greater the overall quality and soundness of the departmental examination interview processes.

Correlational analysis of the data was conducted to determine if there were any relationships between the quality of decentralized examination interview processes

and departmental staffing activities, as well as departmental testing practices. Table 29 presents the correlation matrix based on the data collected via this examination review process and general demographic data for the 16 departments that conducted the 64 examination interview processes reviewed. The major findings based upon correlational analysis can be summarized as follows:

- Interview development methodology correlates significantly with staff participation in the SATP program and the appropriateness with which departments utilize scoring models.
- Interview content correlates significantly with staff participation in the SATP program and departmental use of job analysis.
- Interview question construction correlates significantly with staff participation in the SATP program, departmental use of job analysis, and the appropriateness with which departments utilize scoring models.
- Scoring criteria correlates significantly with staff participation in the SATP program, departmental use of job analysis, and the appropriateness with which departments utilize scoring models.
- Pass point methodology correlates significantly with staff participation in the SATP program, departmental use of job analysis, the appropriateness with which departments utilize scoring models, and the use of structured scoring criteria.

### Performance Test Development

Fifteen (17%) of the 86 examination processes reviewed incorporated the use of a performance test. The performance test component was either utilized as the sole selection instrument, one of multiple instruments comprising the examination process, or in conjunction with an examination interview process as a pre-interview exercise. The 15 performance tests reviewed were conducted by eight departments. Four large departments conducted four (27%) of the performance tests reviewed; three medium departments conducted six (40%); and, one small department conducted five (33%).

A number of factors related to the development of the performance tests were evaluated for the 15 testing processes, including the following:

- The methodology followed in developing the performance test

**Table 29**  
**Correlations among Staffing and the Quality of Interview Processes**

	n	Correlations													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Department size	86														
2. Size of examination staff	80	.48**													
3. Training through SATP	86	.31**	.41**												
4. I/O training/education	80	.48**	.99**	.43**											
5. Professional organizations	80	-.41**	.11	.21	.08										
6. Job analysis	86	.17	.35**	.35**	.33**	.43**									
7. Instrument selection	85	.10	.17	.19	.17	.31**	.41**								
8. Scoring model	86	.05	.14	.20	.15	.28**	.39**	.57**							
9. Development methodology	64	.33**	.14	.34**	.16	.07	.21	.26*	.32**						
10. Interview content	64	.13	.13	.41**	.16	.06	.29**	.21	.23	.17					
11. Question construction	64	.32**	.31**	.30**	.35**	.10	.29**	.41**	.26**	.49**	.30**				
12. Scoring criteria	64	.01	.05	.52**	.07	.43**	.36**	.31**	.37**	.24	.30**	.35**			
13. Pass point methodology	61	-.07	.05	.50**	.04	.41**	.33**	.17	.40**	.15	.31**	.18	.84**		
14. Rater training/preparation	63	.08	.06	.10	.09	.20	.25*	.29*	.27*	.43**	.31*	.51*	.11	.12	
15. Application of scoring criteria	62	.05	.13	.31*	.17	.19	.20	.15	.28*	.14	.31*	.35**	.55**	.59**	.27*

*Note:* For department size, 1 = small department, 2 = medium department, 3 = large department. For training through SATP (SPB's Selection Analyst Training Program), 0 = no staff attendance in SATP classes, 1 = some or all staff have attended some or all classes, but no one is certified, 2 = some staff SATP-certified, 3 = all staff SATP-certified.

\*\* - < .01. \* - < .05.

- The content of the instrument based upon the KSAs identified for assessment
- The construction of the exercises or simulated job tasks;
- The scoring criteria
- The pass point setting methodology
- The use of SMEs in the development of the instrument

Fourteen (93%) of the 15 performance tests reviewed met or exceed the minimum standards for test development. The one performance test which failed to meet the minimum test development standards lacked sufficient documentation pertaining to scoring criteria and pass point methodology. All of the tests were deemed to be job-related, soundly developed techniques for appropriately assessing candidate

qualifications. The one test that lacked sufficient scoring criteria ultimately had been developed soundly; it simply lacked an adequate scoring system. It should be noted that this test, while administered to candidates, ultimately was not used in determining candidate performance in the overall examination process.

Table 30 illustrates the evaluation results in the Performance Test Development category.

**Table 30**  
***Evaluation Results – Performance Test Development***

Rating Dimension	N	Mean	Median	Frequency Results	
				Rating	Frequency
Development methodology	15	2.13	2.00	3	2 (13%)
				2	13 (87%)
				1	0
				0	0
SME participation	14	2.00	2.00	2	14 (100%)
				1	0
				0	0
Instrument content	15	2.47	2.00	3	7 (47%)
				2	8 (53%)
				1	0
				0	0
Instrument Construction	15	2.73	3.00	3	11 (73%)
				2	4 (27%)
				1	0
				0	0
Scoring Criteria	15	2.73	3.00	3	12 (80%)
				2	2 (13%)
				1	1 (7%)
				0	0
Preliminary pass point setting activities, as appropriate	3	2.00	2.00	3	1 (33%)
				2	1 (33%)
				1	1 (33%)
				0	0
Pass point setting methodology	14	2.57	3.00	3	9 (60%)
				2	4 (27%)
				1	1 (7%)
				0	0



### Performance Test Administration

In addition to evaluating the development of the 15 performance tests, various administration activities of those instruments were reviewed, including the following:

- The instructions provided to candidates
- The extent to which raters were trained/prepared for their assessor roles
- The qualifications of the raters
- The application of the scoring criteria by the raters in rating candidate performance in the performance tests
- The availability of examination statistics
- The extent to which the administration of the performance tests resulted in adverse impact

Review of the instructions provided to candidates indicated that 10 (67%) of the 15 performance tests had provided clearly worded, standardized instructions to the candidates. Two (13%) of the performance tests failed to provide adequate instructions to the candidates, although the instructions that were provided attempted to provide the candidates with testing expectations. Candidate instructions for the remaining three (20%) performance tests were not available for review; therefore, the presence and/or adequacy of such instructions could not be determined.

Based on available records, 11 (73%) of the 15 performance tests provided for sufficient training and preparation of the raters prior to administration of the performance tests. Documentation regarding rater training was not available for one of the performance tests reviewed; therefore, the presence and/or adequacy of such training could not be determined. Two (13%) performance tests utilized scannable, objectively scored answer sheets; therefore, the use of raters was not required. The qualifications of the raters used to score the performance tests was also a facet of the evaluation process. Documentation of rater qualifications for 12 (92%) of the 13 performance tests that utilized raters indicated that the raters were appropriately qualified to evaluate candidate qualifications. For the other performance test process (8%) which utilized raters, documentation of rater qualifications was not sufficient to determine the appropriateness of the raters.

To determine the extent to which the scoring criteria or established scoring process had been applied by the raters, a review of rating sheets, rater notes, final scoring documentation, and actual scoring criteria was completed. This review focused on

the application of the performance test scoring criteria as opposed to the quality of the scoring criteria. A comparison was made among the various records and documentation to determine the extent to which the performance tests had been scored in a consistent manner. Documentation and records for the 13 performance tests which were administered as stand-alone processes were reviewed. Of the 13 performance tests reviewed, documentation for all of the processes indicated consistent scoring practices by the raters. No scoring inconsistencies were noted in any of the 13 performance tests.

The only deficiency noted on a wide-scale basis for the performance tests reviewed was the absence of examination statistics and corresponding reports. The specific reports identified in the review criteria included examination control documentation and bottom line hiring reports. In reviewing the files and documentation for the 13 performance tests which were administered as stand-alone processes, it was determined that the specified reports were available for only five (38%) of the 13 tests. For one performance test (8%), only partial documentation was available. For the majority of other performance tests conducted, the departments which had conducted the performance tests do not currently use the SPB's on-line examination scoring system, from which these reports are typically generated.

Table 31 illustrates the evaluation results in the Performance Test Administration category.

Adverse impact data were available for six of the 15 performance tests reviewed; documentation/data were not available for the other nine performance tests. Of the six performance tests for which documentation/data were available, two (33%) of the six tests resulted in adverse impact against one or more candidate groups. Neither of the performance tests had been validated or developed based upon a job analysis, and, therefore, conclusions about whether the content of the tests was related to business necessity could not be determined.

#### Supplemental Application Development and Administration

Two of the 86 examination processes reviewed incorporated the use of a supplemental application. The supplemental applications both utilized the behavioral consistency method of acquiring information and data regarding candidates' backgrounds, expertise, training, and experience. Both of the supplemental applications functioned as one of multiple instruments comprising the examination process. The two supplemental applications were utilized by one medium and one large department.

**Table 31**  
***Evaluation Results – Performance Test Administration***

Rating Dimension	n	Mean	Median	Frequency Results	
				Rating	Frequency
Candidate instructions	12	2.67	3.00	3	10 (83%)
				2	0
				1	2 (17%)
				0	0
Rater training/preparation	12	2.75	3.00	3	10 (83%)
				2	1 (8%)
				1	1 (8%)
				0	0
Rater qualifications	12	2.00	2.00	2	12 (100%)
				1	0
				0	0
Application of scoring criteria	13	2.00	2.00	2	13 (100%)
				1	0
				0	0
Availability of examination statistics and corresponding reports	8	1.38	2.00	2	5 (63%)
				1	1 (13%)
				0	2 (25%)

A number of factors related to the development of the supplemental applications were evaluated for the two supplemental applications, including the following:

- The content of the application based upon the KSAs identified for assessment
- The construction of the application
- The scoring criteria
- The pass point setting methodology
- The use of SMEs in the development of the instrument.

Various factors related to the administration or use of the supplemental applications were reviewed, including the following:

- The instructions provided to candidates

- The extent to which raters were trained/prepared for their assessor roles
- The qualifications of the raters
- The application of the scoring criteria by the raters in rating candidate qualifications
- The availability of examination statistics
- The extent to which the administration of the supplemental applications resulted in adverse impact

In reviewing the two supplemental applications, it was determined that both of the instruments met or exceeded minimal criteria in all of the development and administration dimensions reviewed with the following two exceptions. Scoring criteria for one of the supplemental applications was not available for review; therefore, it could not be determined if the evaluation of candidate qualifications was based on objective, standardized, job-related criteria. The other supplemental application lacked sufficient examination control and bottom line hiring documentation.

Table 32 and Table 33 illustrate the evaluation results in the Supplemental Application Development and Administration categories, respectively.

A review of adverse impact data revealed that one of the supplemental application processes (50%) resulted in adverse impact against one or more candidate groups. The supplemental application had not been validated nor developed based upon a job analysis, and, therefore, conclusions about whether the content of the supplemental application was related to business necessity could not be determined.

#### Use of Ratings of Education, Training and Experience

Nine (10%) of the 86 examination processes reviewed consisted solely of ratings of candidate education, training, and experience (E&Es). Of the nine E&E processes reviewed, two (22%) were conducted by large departments, two (22%) by medium departments, and five (56%) by small departments. The E&Es simply provided for a structured rating of candidate qualifications as presented on the standard State application. Therefore, an actual testing instrument was not developed. The only development activity occurred in the preparation of scoring criteria with which to rate candidate applications. Thus, the following factors were considered in the review of the E&Es:

**Table 32**  
***Evaluation Results – Supplemental Application Development***

Rating Dimension	n	Mean	Median	Frequency Results	
				Rating	Frequency
Development methodology	2	2.00	--	3	0
				2	2 (100%)
				1	0
				0	0
SME participation	2	2.00	--	2	2 (100%)
				1	0
				0	0
Instrument content	2	2.50		3	1 (50%)
				2	1 (50%)
				1	0
				0	0
Question construction	2	3.00	--	3	2 (100%)
				2	0
				1	0
				0	0
Scoring Criteria	1	3.00	--	3	1 (100%)
				2	0
				1	0
				0	0
Pass point setting methodology	1	3.00	--	3	1 (100%)
				2	0
				1	0
				0	0

- The scoring criteria developed
- The qualifications of those individuals rating the candidate applications
- The application of the scoring criteria
- The availability of examination statistics
- The extent to which the use of the E&Es resulted in adverse impact

In reviewing the nine E&E processes, it was determined that serious deficiencies existed in the scoring criteria utilized for many of the processes. Scoring criteria

**Table 33**  
***Evaluation Results – Supplemental Application Administration***

Rating Dimension	n	Mean	Median	Frequency Results	
				Rating	Frequency
Candidate instructions	2	3.00	--	3	2 (100%)
				2	0
				1	0
				0	0
Rater training/preparation	2	2.75	--	3	1 (50%)
				2	1 (50%)
				1	0
				0	0
Rater qualifications	2	2.00	--	2	2 (100%)
				1	0
				0	0
Application of scoring criteria	2	2.00	--	2	2 (100%)
				1	0
				0	0
Availability of examination statistics and corresponding reports	2	1.38	--	2	1 (50%)
				1	1 (50%)
				0	0

were available for six (67%) of the processes, and the criteria for all six processes were insufficient to determine candidate qualifications and differentiate among levels or degrees of candidate expertise. The criteria were developed to simply assign higher scores to those candidates with a greater quantity of experience or education without regard for the quality of that experience. Thus, candidates with comparable levels of expertise, but differing lengths of experience would not receive comparable scores. There was no job-related rationale upon which the length of experience or education parameters were based; they were simply a result of arbitrarily assigning greater value to longer periods of time spent on a job without regard to the expertise actually possessed.

To determine the extent to which the scoring criteria had been applied by the raters, a review of candidate applications, final scoring documentation, and the actual scoring criteria was completed. This review focused on the application of the E&E scoring criteria as opposed to the quality of the scoring criteria. A comparison was made among the various records and documentation to determine the extent to

which the E&Es had been scored in a consistent manner. Of the six E&Es reviewed, documentation for five (83%) of the E&E processes indicated consistent scoring practices by the raters. In addition, based on available records, five (56%) of the nine E&E processes provided for the use of raters who were appropriately qualified to evaluate candidate qualifications.

Table 34 illustrates the evaluation results for the E&E processes.

**Table 34**  
***Evaluation Results – E&E Processes***

Rating Dimension	n	Mean	Median	Frequency Results	
				Rating	Frequency
Scoring Criteria	6	.67	1.00	3	0
				2	0
				1	4 (67%)
				0	2 (33%)
Rater qualifications	5	2.00	--	2	5 (100%)
				1	0
				0	0
Application of scoring criteria	6	1.67	2.00	2	5 (83%)
				1	0
				0	1 (17%)

A review of adverse impact data found that none of the E&E processes resulted in adverse impact.

### Bottom Line Results

The final evaluation conducted for each of the 86 examination processes reviewed was an assessment to determine the extent to which these processes may have resulted in bottom line adverse impact against any candidate group. Bottom line hiring data and documentation were available for 75 (87%) of the 86 examination processes reviewed. In reviewing available bottom line data and documentation for these 75 examination processes, it was determined that administration of 15 (20%) of the examination processes resulted in adverse impact against one or more candidate groups. Of those 15 occurrences of adverse impact, seven occurrences (47%) could be defended on the basis of business necessity, as those seven examinations had validation evidence documenting their job-relatedness. The

examination processes resulting in the other eight occurrences of adverse impact had not been validated nor developed based upon job analysis, and, therefore, conclusions about whether the content of the examination processes was related to business necessity could not be determined.

## **Cost Comparison**

A cost comparison between centralized and decentralized testing was conducted by project staff in May and June 2002. The purpose of the cost comparison was to determine the extent to which decentralized testing efforts may be more costly than centralized testing efforts. In conducting the preliminary research for the cost comparison and designing the methodology for the collection and analysis of the data, project staff determined that departments conducting examination processes did not routinely track actual costs incurred in the development and administration of the examination processes. To ensure that suitable data were collected, project staff determined that the cost comparison would focus on a single servicewide job classification and that project staff would work closely with departmental staff to collect cost data reflective of any known costs, as well as estimated and projected costs.

In determining the job classification to be the subject of the cost comparison, it was noted that the SPB had conducted a centralized open examination process for the servicewide classification of Office Technician (Typing) in the Spring of 2002. Due to the availability of such timely data, the Office Technician (Typing) classification was selected as the focus of the cost comparison. Data from the SPB's 2002 servicewide open examination process for Office Technician (Typing) were collected to represent cost data for centralized testing. Data from three separate departmental examination processes for the Office Technician (Typing) classification were collected to represent the cost data for decentralized testing.

The three departmental examination processes represented a sample of the departmental examination processes conducted for the Office Technician (Typing) job classification from late 2000 through Spring 2002. In determining the departmental examination processes to be included in this cost comparison, project staff sought to identify departmental examination processes similar to the SPB's examination process in terms of examination base (i.e., an open rather than promotional process), examination plan (i.e., an examination process consisting of the use of a written examination rather than a different type of instrument), and size of applicant/candidate group. The three departments represented in the cost



comparison were the California Highway Patrol (CHP), the Department of Motor Vehicles (DMV), and the Department of the Youth Authority (CYA).

The examination plans for the examination processes comprising the cost comparison were similar, although not exact. The SPB's examination process was comprised of a newly developed written examination, weighted 100%, as the sole selection component. All three of the departmental examination processes included the use of a written examination; however, the written examinations used were instruments that had been previously developed and administered and were simply reviewed and revised for the current examination processes. In addition, two of the departmental examination processes included an examination interview in addition to the written examination. Table 35 depicts the selection components, as well as the testing bases and number of applicants, for the examination processes reviewed.

**Table 35**  
***Office Technician (Typing) Examination Processes***  
***Reviewed for Cost Comparison***

<b>Department</b>	<b>Selection Components</b>	<b>Testing Base</b>	<b>Number of Applicants</b>
State Personnel Board	Written Examination	Open	11,544
California Highway Patrol	Written Examination and Examination Interview	Promotional	250
Department of Motor Vehicles	Written Examination and Examination Interview	Promotional	54
Department of the Youth Authority	Written Examination	Promotional	70

In collecting the data for the cost comparison, project staff worked with SPB staff and departmental staff to collect costs associated with the respective examination processes. A systematic approach was used to collect the cost data to ensure that the data reflected accurate and comparable cost figures and estimates. It should be noted that while none of the three departments tracked actual examination process costs, the data collected reflected some known costs, as well as estimated and projected costs derived from departmental staff in conjunction with project staff queries. Data collection efforts focused on hard dollar costs, staff and resources costs, and timeline requirements, including the following:

- Staff time required to plan, develop, announce, schedule, and administer the examination process, including analyst, technician, and proctor requirements
- Staff time required to screen candidate applications
- Staff time required to prepare and distribute candidate notices
- Costs associated with the scanning of candidate answer sheets and the scoring of candidate test performance
- Facility costs for examination administration
- Supply and printing costs for the preparation of examination booklets, examination bulletins, and other testing materials

Based upon the data collected, costs were calculated initially for each of the three departmental examination processes and the SPB's examination process. The cost calculations were based on the assumption that a written examination was the sole selection component in the examination process. Using the total costs calculated for each of the examination processes, costs per applicant, candidate, and eligible were calculated. Table 36 illustrates the costs calculated for each of the examination processes. The complete cost data initially calculated for the SPB's examination process and the three departmental examination processes is provided in *Appendix K*.

The cost data were further analyzed and interpreted to produce cost estimates for what would be similarly sized applicant/candidate groups participating in like examination processes (i.e., an examination process consisting of a newly developed written examination as the sole selection component). The departmental data were augmented using elements of the SPB's cost data to generate projected cost estimates to conduct an examination process comprised of a newly developed written examination as the sole selection component for an applicant/candidate group of a size comparable to that having participated in the SPB's examination process. This interpretation of the data represented the projected cost differences between a centralized testing effort and a representative sample of decentralized testing efforts. As analyzed, the data indicated that the projected costs to develop and conduct centralized and decentralized examination processes of similar examination type with similarly sized applicant/candidate groups would be comparable. Table 37 illustrates the projected costs differences between centralized and decentralized testing efforts.

**Table 36**  
**Office Technician (Typing) Examination Processes Costs**

<b>Department</b>	<b>Total Cost for Examination Process</b>	<b>Cost per Applicant <sup>1</sup></b>	<b>Cost per Candidate <sup>2</sup></b>	<b>Cost per Eligible <sup>3</sup></b>
State Personnel Board	\$138,646.00	\$12.01 (n = 11,544)	\$16.82 (n = 8,242)	\$28.61 (n = 4,846)
California Highway Patrol	\$ 5,759.19	\$23.04 (n = 250)	\$24.00 (n = 240)	\$50.97 (n = 113)
Department of Motor Vehicles	\$ 2,439.61	\$45.18 (n = 54)	\$59.50 (n = 41)	\$87.13 (n = 28)
Department of the Youth Authority	\$ 3,412.70	\$48.75 (n = 70)	\$56.88 (n = 60)	\$97.51 (n = 35)

**Notes:**

<sup>1</sup> Applicant is defined as an individual submitting an application to participate in the examination process

<sup>2</sup> Candidate is defined as an applicant meeting the minimum qualification requirements and invited to participate in the written examination

<sup>3</sup> Eligible is defined as an individual passing the written examination and whose name was (or would have been) placed on the eligible list

Further analysis of the projected cost data for centralized and decentralized testing efforts indicated a potential difference in costs. It was determined that such cost differences could occur in those instances where an applicant would be required to compete in multiple decentralized departmental examination processes in order to establish list eligibility in more than one department rather than competing in a single centralized examination process that would result in list eligibility on a single servicewide eligible list. The resulting servicewide eligible list would provide the applicant/candidate the equivalent of list eligibility in multiple departments. In this instance, a single centralized examination process would be less expensive and more efficient from the standpoint of State resources, as well as the applicant/candidate's time and effort, than would multiple decentralized departmental processes. Based upon the data, the costs to conduct multiple decentralized examination processes rather than a single centralized examination process were projected to be nearly three times greater for the decentralized testing efforts. These data as representative of the cost of testing duplicity would serve as a baseline value for further cost estimates. Table 38 illustrates the increased costs associated with duplicative decentralized testing processes as derived from the data in the cost comparison herein.

**Table 37**  
**Projected Costs**  
**for Centralized and Decentralized Examination Processes**

<b>Examination Process Type</b>	<b>Projected Cost for Examination Process <sup>1</sup></b>	<b>Cost per Applicant (n = 11,544)</b>	<b>Cost per Candidate (n = 8,242)</b>	<b>Cost per Eligible (n = 4,846)</b>
Centralized Testing Process (costs derived from State Personnel Board data)	\$138,646.00	\$12.01	\$16.82	\$28.61
Decentralized Testing Process (costs derived from CHP data)	\$129,934.80	\$11.26	\$15.76	\$26.81
Decentralized Testing Process (costs derived from DMV data)	\$128,973.81	\$11.17	\$15.65	\$26.61
Decentralized Testing Process (costs derived from CYA data)	\$129,673.80	\$11.23	\$15.73	\$26.76

Note:

<sup>1</sup> Examination process is defined an examination process comprised of a newly developed written examination as the sole selection component for a large applicant/candidate group representative of the State's large applicant/candidate groups

**Table 38**  
**Projected Duplicative Costs for Decentralized Examination Processes**

<b>Examination Process Type</b>	<b>Projected Cost for Examination Process <sup>1</sup></b>	<b>Cost per Applicant (n = 11,544)</b>	<b>Cost per Candidate (n = 8,242)</b>	<b>Cost per Eligible (n = 4,846)</b>
Centralized Testing Process (costs derived from State Personnel Board data)	\$138,646.00	\$12.01	\$16.82	\$28.61
Multiple Decentralized Testing Processes (costs derived from summing CHP, DMV, and CYA data)	\$388,582.41	\$33.66	\$47.15	\$80.19
<i>Increased Costs for Decentralized Testing Processes</i>	<i>\$249,936.41</i>	<i>\$21.65</i>	<i>\$30.33</i>	<i>\$51.58</i>

Note:

<sup>1</sup> Examination process is defined an examination process comprised of a newly developed written examination as the sole selection component for a large applicant/candidate group representative of the State's large applicant/candidate groups

## **State Personnel Board Hearing**

A hearing before the State Personnel Board was held July 9, 2002. The hearing solicited comments regarding the preliminary findings and recommendations issued for this project in June 2002. In addition, the hearing functioned as the final data collection element for this project. Public comment focusing on issues relevant to the evaluation of the State's decentralized testing program was solicited via notice of the hearing which was disseminated to all State agencies, departments, and employee organizations. Project staff also invited specific stakeholder groups to offer comments at the hearing.

The format for the hearing followed a schedule established as a result of specific stakeholders indicating an interest in addressing the Board. To begin the hearing, project staff provided a brief overview of the project to review the State's decentralized testing program, as well as discussed the cost data collected and preliminary findings of the cost comparison. Following the introductory comments by project staff, comments were provided by 14 individuals representing a variety of interests and groups, including departments, unions, the assessment profession, as well as individual employee/candidate interests.

The comments provided by those addressing the Board covered a range of issues and topics. It should be noted that the integrity of the project methodology and the data collection and analysis efforts to date were not challenged by any of the speakers. In fact, the majority of speakers recognized the Board's support of this project, as well as SPB staff efforts thus far with the project. Following is a summary of the comments provided at the hearing presented by stakeholder interests:

### Departmental representatives

The comments provided by the departmental representatives, including the SPIN, Examination Supervisors' Forum, Health Services, and Corrections representatives, expressed a general recognition that the State's decentralized testing program could be improved, although concern that enhanced selection requirements, as outlined in the preliminary recommendations, would require increased departmental resources, inclusive of both additional staff resources and dollar resources, was also expressed. It was the general sentiment of the departmental representatives that such increases in staffing and resources would not be likely given the State's current budget challenges which could be a further detriment to the testing program if departments were to be held to higher standards with no increase in

staffing and resources. It was further stated that most departments were not currently staffed nor prepared to conduct sound examination processes, per recognized legal and professional standards and principles, as recommended in the preliminary report. The departmental representatives indicated that while they appreciated the SPB striving to improve the State's testing program, they believed that the preliminary recommendations as drafted were calling for an optimal testing program rather than a testing program designed to meet minimum legal and professional rigor. Comments from the departmental representatives also addressed the following topics/issues:

- The Selection Analyst Training Program received generally positive comments, with departments supporting SPB efforts toward training and developing existing departmental examination staff, although it was suggested the training curriculum be reviewed to ensure that it represents the SPB's current standards and direction for the State's testing program.
- The SPB's current *Selection Manual* was mentioned repeatedly as needing to be revised to document the testing standards and requirements established by the SPB for the State's testing program.
- The SPB's on-line applicant tracking/examination scoring system was identified as needing to be modified and updated to provide adequate support for a variety of selection processes of varying degrees of complexity, such as the types of examination processes discussed in the Selection Analyst Training Program.
- It was repeatedly mentioned that the SPB needs to provide consistent information and interpretation of laws, rules, policies, and practices to its stakeholders, regardless of which division(s) and/or departmental staff are providing the information.

### Union representatives

The union representatives, including the Association of California State Supervisors, the California Association of Psychiatric Technicians, the California Department of Forestry Firefighters Association, and the California State Employees' Association, indicated support for increased SPB oversight and control in the State's testing program in

line with the preliminary findings and recommendations issued for the project. One of the unions went so far as to call for the complete recentralization of the State's testing program to the SPB. The majority of union representatives also indicated that the decentralized testing program as currently administered by departments has been weakened by real and perceived instances of favoritism, bias, and discrimination. There was a general sentiment expressed that many of the decentralized examination processes conducted by departments are not measures of candidate job-related qualifications and, therefore, do not result in appointments made on the basis of a job-related rationale.

The preliminary findings and recommendations were not challenged with regards to how they could not or would not meet the unions' concerns for fair, unbiased testing but rather reinforced by the comments of a number of the union representatives.

#### Assessment profession representative

The County of Sacramento's Personnel Services Division Chief provided expert comments regarding how the preliminary findings and recommendations issued for this project compared to the current state of the assessment profession. In the speaker's expert opinion the draft recommendations were interpreted as calling for testing standards in line with minimal legal and professional standards and principles, rather than standards to establish an optimal testing program. The speaker further indicated support for the SPB's Selection Analyst Training Program in terms of its content and curricula, indicating that over the past few years she has sent a number of her examination staff to a number of the classes offered.

#### Additional hearing participants

Three individuals representing themselves as stakeholders in the State's testing process provided comments during the hearing. The general sentiments expressed by these individuals supported the preliminary findings and recommendations issued for this project, with the majority of the individuals calling for the recentralization of the State's testing program to ensure merit and fairness in the selection process. A number of comments made by these individuals addressed a lack, or perceived lack of, fairness and objectivity in the State's selection process. The individuals cited examples of favoritism



contributing to results of examination processes, as well as being the basis for appointment decisions within the State's civil service.

These individuals also addressed issues outside of the scope of this study, focusing on perceived inadequacies with the hiring interview/appointment stage of the hiring process, as well as the shortcomings of the State's appeals process in resolving problems within the selection process.

### Summary analysis

In further interpreting the comments provided in the hearing, a common theme was noted in the comments of one departmental speaker and one union speaker in terms of the value and benefits of increasing the soundness of testing efforts. The departmental representative indicated that the recent use of better, more sound examination processes had resulted in higher quality candidates being placed on the eligible lists. The speaker indicated that while the better, more sound examination processes had resulted in higher candidate fail rates in the examination processes, a greater number of appeals, and fewer candidates ultimately being placed on the eligible lists, the more sound examination processes had resulted in the identification of better qualified candidates. In conjunction with these remarks, a representative of one of the unions indicated that he had seen first-hand a decline in the quality of new hires appointed as a result of the use of less sound, less rigorous testing procedures in a specific department. The comments of both of these speakers spoke directly to the utility of sound testing and the appropriateness of the preliminary findings and recommendations issued for this project.

### Written comments

In addition to the public hearing, public comment in the form of written comments was also solicited. Several stakeholders provided written comments regarding the overall status of the State's decentralized testing program, as well as providing comments directly related to the preliminary findings and recommendations of the project. Written comments were provided by the Ad Hoc Committee on Personnel, representatives of three departments, an individual employed by the Department of Corrections, and representatives from two unions. The written comments provided by the two unions were written synopses of oral comments provided by their respective representatives at the hearing.

The written comments submitted by the seven stakeholders covered a range of issues and topics. The majority of stakeholders noted the Board's support of this



project, as well as SPB staff efforts. In addition, the project methodology and the data collection and analysis efforts to date were not at issue in the written comments. The stakeholder comments primarily addressed the impact of the preliminary recommendations issued in conjunction with this project, as well as implementation challenges that would result from Board-approval of some of the preliminary recommendations.

The written comments provided by the departmental representatives, including the Ad Hoc Committee on Personnel, the Franchise Tax Board, and the Departments of Justice and Mental Health, supported the State Personnel Board's review of the decentralized testing program and expressed general agreement with the preliminary recommendations. The comments indicated a unanimous commitment to improving the State's testing program and support for the SPB's efforts to establish uniform standards for the testing program; however, the comments further indicated that additional staff and budget resources would be required by most departments in order to effectively implement the majority of the preliminary recommendations aimed at improving the testing program. Comments from all of the departmental representatives addressed the following topics/issues:

- The Selection Analyst Training Program received positive comments, with departments supporting SPB efforts toward training and developing existing departmental examination staff. It was suggested that the training curriculum be reviewed to ensure that it represents the SPB's current standards and direction for the State's testing program and that it results in the effective training of course participants.
- The SPB's current *Selection Manual* was noted as needing to be revised to document the testing standards and requirements established by the SPB for the State's testing program.

The comments provided by the individual employed by the Department of Corrections detailed an examination interview process in which he had recently participated. The individual indicated that due to short-comings in the interview process, he was unsuccessful in the examination process while other candidates with lesser qualifications and/or closer working relationships with the interview panel members had been successful in the interview process. His assertion was that recent departmental promotional testing had not resulted in the accurate assessment of candidate qualifications.

A listing of the groups providing oral testimony during the hearing, as well as those

groups providing written testimony in conjunction with the hearing, is provided in *Appendix H*.

## **Summary of Findings and Discussion**

The following section of this report discusses in detail the results and findings of the data collection activities documented herein.

Based upon thorough analysis of the data collected for this project, numerous findings related to the merit and efficiency of the State's decentralized testing program have emerged. While the data collection elements were pursued as independent variables, many common findings were obtained from multiple data sources. A number of efficiencies were cited for the State's decentralized testing program, with the most frequently cited being the frequency and timeliness with which departments can conduct decentralized testing processes, the geographically localized testing efforts that departments can conduct to meet specific departmental selection needs, and the familiarity of departmental examination staff with specific departmental selection needs. While a number of efficiencies were noted in the decentralized testing program, a number of issues and concerns were also noted by many of the program stakeholders interviewed, as well as being manifest in the data collected by project staff. In considering the thematic content of the issues and concerns raised, a number of common deficiencies with the program were identified through multiple data sources. The most commonly cited deficiencies were related to the quality of the testing processes conducted under the decentralized testing program.

Although a number of issues surfaced with regards to the quality and soundness of decentralized testing practices and processes, there was evidence that in some instances departments are conducting sound testing processes, and in fact, many departments and agencies are actively seeking to improve their current testing practices and processes. However, the magnitude of the deficiencies noted in many departmental testing processes and practices jeopardizes the overall value of the State's testing program. For the most part, the deficiencies which were manifest in the data involve basic elements of integral importance to the overall soundness of an examination process. Such deficiencies tend to overshadow other aspects of an examination process which might be conducted in a satisfactory manner. For example, in some instances the data indicated that an examination process was comprised of acceptable selection instruments; however, there were no data documenting the job-relatedness of those instruments. Thus, while the examination plan appeared to be appropriate, the job-relatedness of the examination process

was at question. It should be noted, though, that many of the deficiencies are not necessarily a result of decentralized testing but, rather, a result of poor testing practices, untrained or inexperienced staff, and/or a lack of value placed on the role of testing by departmental leadership.

An analysis of the data indicated that those departments which have most prevalently utilized the SPB's Selection Analyst Training Program for examination staff training conduct higher quality, better job-related testing processes than do those departments whose staff have received less training or no training through the program. Data collected indicates that large departments and small departments have a greater proportion of examination staff who have completed the training program curriculum. The primary reasons cited by departments for the lack of staff participation in the training program include insufficient financial resources and the availability of staff to attend due to workload issues.

In further analyzing the data collected, several program deficiencies were noted as a result of the review of actual decentralized testing processes. In general, the decentralized testing processes reviewed fail to use highly structured and standardized assessment techniques. The most serious deficiency identified is the lack of job-relatedness of many of the testing processes. A large majority of decentralized testing processes are not based on job analyses which meet the job analytic requirements set forth by the *Uniform Guidelines*, nor do these processes demonstrate evidence of validity (i.e., documented evidence of a relationship between test performance and job performance). The absence of sufficient job analytic data and corresponding validation evidence is observed in all types of testing components, including written examinations, examination interviews, performance tests, and application screening processes. This tends to be the case regardless of department size. In addition, administration of a number of the testing processes resulted in adverse impact against one or more candidate groups. In reviewing available data, it was noted that administration of 15 of the processes resulted in adverse impact. Of those 15 occurrences of adverse impact, seven occurrences (47%) could be defended on the basis of business necessity, as those seven examinations had validation evidence documenting their job-relatedness. The examination processes in the other eight occurrences had not been validated nor developed based upon job analysis, and, therefore, conclusions about whether the content of the examination processes was related to business necessity could not be determined.

Another noted deficiency is the lack of job-related pass points used for written examinations and examination interview processes. The majority of pass points reviewed fail to demonstrate that the required level of competency was related to acceptable job performance. The failure to utilize job-related pass points was

observed in departments of all sizes, although large departments are more likely to rely on job-related pass points for written examinations than are medium or small departments. In addition to noted concerns regarding the job-relatedness of pass points utilized for examination interview processes, numerous deficiencies related to the scoring criteria used for the examination interview processes were also identified. The examination interview process comprises more than 50 percent of the examining that occurs in the State's selection system; thus, deficiencies in the examination interview process have the potential to significantly affect the integrity of the State's selection system. The documented deficiencies occurred in a number of examination interview processes regardless of department size, examination base (e.g., open, promotional), and/or job level.

Further analysis of the data indicated that these deficiencies represent failings of decentralized testing processes to adhere to recognized assessment standards and principles, including the *Uniform Guidelines*, the Civil Rights Acts of 1964 and 1991, the Americans with Disabilities Act, the *Principles for the Validation and Use of Personnel Selection Procedures*, and the *Standards for Educational and Psychological Testing*, as well as professional literature. Of further note is the lack of clarity in current State law, rules, and regulations pertaining to the specific design and development of employment testing processes and individual selection instruments. Thus, while legal and professional standards and principles may delineate sound selection, greater specificity in State law, rules, and regulations would help to assure the soundness of the State's testing program.

The union and employee advocacy group stakeholders cited the following as the primary program deficiencies:

- Favoritism and pre-selection in hires and appointments
- Favoritism and pre-selection in examination processes
- Lack of structure in examination interviews (i.e., interviews administered as a component in the civil service examination process as opposed to hiring interviews), including the use of non-job-related questions in the interview, the reliance on subjective, non-job-related scoring criteria, the use of non-expert interview panel members, and the use of panel members familiar with candidates
- Lack of testing expertise of some departmental examination staff
- Lack of adequate communication with candidates, including a lack of notification of examination activities, failure to provide feedback to candidates participating in examination processes, and failure to notify

candidates when they are reachable on eligible lists

Many of the above-noted deficiencies are consistent with the issues and deficiencies noted by the United States Merit Systems Protection Board in its reviews of the Federal government's decentralized testing program, as well as those noted by several of the State government representatives contacted for data collection purposes. Additionally, data collected by means of the departmental questionnaire, the departmental interviews, and the review of the decentralized testing processes revealed similar findings.

An additional issue which was raised by the union and employee advocacy group stakeholders was the perception that candidates may not exercise their right to appeal examinations for fear of reprisal. No additional data corroborating this sentiment emerged via any of the other data collection elements, and the large number of appeals filed would tend to disprove this perception. However, without further study, the issue of reprisal or retaliation can not adequately be addressed.

As a final point, a recurrent complaint of the decentralized testing program was directed at the quality of support and consultation provided by the SPB. The complaints focused on the frequency with which conflicting information is provided by SPB staff in response to queries from various stakeholders, including departmental examination staff, State employees, examination candidates, and union/employee advocacy group representatives. Data collected regarding this issue indicates that the conflicting information frequently requires departments to defend their actions to one SPB function based upon direction provided by another SPB function. It was stated by numerous stakeholders that in its role as overseer of the decentralized testing program, the SPB needs to provide consistent interpretation of policies, rules, and regulations, as well as provide departments with a single set of expectations for conducting decentralized testing activities.

Data collection efforts also focused on identifying program improvements. A multitude of suggestions were provided by various program stakeholders. The most commonly cited improvements involve the role of the SPB in the decentralized testing program. Departments called for revisions to the SPB's *Selection Manual* and on-line examination scoring system, as well as greater customer service and support from SPB staff. In addition, union and advocacy groups called for the SPB to play a greater role in overseeing decentralized testing activities.

Based upon data collection and analysis efforts, a series of recommendations for improving the quality and soundness of the State's decentralized testing program have been developed. The recommended program improvements have been

identified as means to improve the State's selection system and provide for increased merit, efficiency, and effectiveness in the testing program. The recommendations include the SPB providing greater oversight and support for the decentralized testing program, as well as identifying specific departmental obligations under decentralization. The recommendations as proposed utilize existing SPB staff and resources. While more comprehensive measures to guide and oversee the decentralized testing program could be proposed, such additional measures would require additional staff and resource allocations for the SPB. Given the State's current budget realities, it is unlikely that the SPB would be allocated additional resources; thus, the proposed recommendations are measures that rely solely on the utilization of existing SPB staff and resources.

The recommendations are presented in the next section of this report.

## Recommendations

The following section of this report outlines staff recommendations based upon the results and findings of the data analysis process documented herein.

An analysis and comprehensive review of the State's Decentralized Testing Program resulted in the emergence of a number of program deficiencies which, if left unchecked, could seriously compromise the quality and effectiveness of the examination processes conducted for the establishment of eligible lists, as well as diminish the integrity of the State's civil service. The major deficiencies identified in the study include the use of examination processes which lack evidence of job-relatedness, a lack of job-related scoring criteria for examination interview processes, and the use of non-job-related pass points for written examinations and examination interview processes. The noted deficiencies encompass a general failure to utilize soundly developed, job-related selection procedures, inclusive of structured, standardized scoring criteria and job-related pass points. The majority of the program deficiencies stem from a lack of staff training and expertise at the department/agency level and a failure of departments to pursue the use of sound, merit-based testing processes on a regular basis, as well as inconsistent direction and support from the SPB.

To mitigate identified program deficiencies and to ensure continued fairness and effectiveness in the State's merit-based selection system and departmental decentralized testing efforts, it is necessary that the SPB provide more proactive, comprehensive oversight and guidance for the State's decentralized testing program. Such increased oversight and guidance will facilitate compliance with applicable federal and state laws and rules pertaining to employment testing and employee selection, as well as further promote the State's merit principle. It is recommended that the increased oversight and guidance provided by the SPB encompass the SPB establishing program standards and guidelines, as well as providing greater expert consultative assistance and hands-on support to departments and agencies. In addition, it is recommended that the SPB assist departments in securing much-needed technical training and expertise in the areas of test construction and psychometrics for their examination staff.

Through a series of short-term and longer-term efforts, it is the intent of the proposed recommendations to resolve noted deficiencies in the decentralized testing program and to ensure that decentralized testing processes meet these criteria:



- The processes be merit-based
- The processes utilize selection instruments that provide for the job-related, objective assessment of candidate qualifications
- The processes adhere to mandated legal requirements
- The processes emulate recognized professional standards and principles for employment testing
- The processes comply with SPB-established program standards and guidelines

Utilizing existing staff and resources, it is proposed that the SPB implement the following recommendations to address the noted deficiencies in the decentralized testing program, as well as foster a more proactive role for the SPB in providing consultative support and guidance related to the testing process to departments and agencies:

- Establish and disseminate specific standards and guidelines for decentralized testing activities
- Offer training to existing departmental examination staff in the areas of job analysis, structured interview development, and job-related pass point setting, which are the main areas of deficiency which have been noted in the decentralized testing program
- Promote departmental recruitment and selection of human resources professionals who possess the requisite testing expertise to function effectively in departmental testing programs
- Provide departmental examination staff with greater access to SPB staff expertise
- Identify resources, both within State service and outside of the State, from which departmental examination staff can access assessment-related guidance and expertise
- Re-evaluate the State's decentralized testing program 24 months following approval of the proposed recommendations to ensure that the program meets established standards and guidelines and provides for job-related, merit-based candidate assessment

It is proposed that implementation of the recommendations be conducted in phases, with the establishment and dissemination of program standards and guidelines occurring first, followed by the offering of training and expertise to departmental examination staff. The following timeline for development and implementation of the proposed recommendations outlines the anticipated short-term and longer-term results:

*Within three months of Board-approval of the recommendations*

The SPB shall develop and disseminate to departments and agencies specific standards for decentralized testing activities to ensure that program expectations are clearly set forth such that departmental examinations programs can be held accountable for compliance.

*Within 12 months of Board-approval of the recommendations*

The SPB shall offer on a no-fee basis training opportunities and assessment-related resources to departments and agencies for departmental use in retooling and/or modifying departmental testing practices to ensure that departmental examination programs adhere to program standards and guidelines.

*Within 24 months of Board-approval of the recommendations*

The major deficiencies in the decentralized testing program will have been addressed by the SPB, and departments and agencies will have been provided with standards, expectations, and means of resolving the program deficiencies.

*Within 30 months of Board-approval of the recommendations*

The SPB will have re-evaluated the decentralized testing program to determine its soundness and to ensure departmental adherence to established program standards and guidelines.

The proposed recommendations encompass the development and dissemination of program standards and guidelines to departments and agencies engaged in decentralized testing efforts, as well as the availability of a series of products and services, to ensure that departments and agencies are aware of their obligations concerning their testing programs. It is also the intent of these recommendations to ensure that departments and agencies are afforded access to necessary training and expert resources to ensure that their examination unit staff (inclusive of existing staff and subsequently hired staff) are prepared to meet established program standards and guidelines.

It is proposed that the following recommendations and corresponding timeline be implemented by the SPB to address noted deficiencies in the decentralized testing program:

Recommendation	Proposed Implementation Timeline
1. The SPB shall promulgate the program standards for the decentralized testing program as outlined in <i>Appendix L</i>	Spring 2003, (pending Board approval of this report)
2. The SPB shall notify departments and agencies engaged in decentralized testing of the requirement that departmental testing efforts comply with program standards, as outlined in <i>Appendix L</i> , within 24 months of Board approval of such program standards	Spring 2003
3. The SPB shall offer to host a monthly (or similarly scheduled) discussion group for Examination Managers/Supervisors to discuss and address selection-related issues and topics; share success stories; and, provide a state-wide forum to resolve problems of a selection/testing nature	Spring 2003
4. The SPB shall offer a <i>hotline</i> support line to address technical examination/selection-related questions and issues	Spring 2003

Recommendation	Proposed Implementation Timeline
<p>5. The SPB shall prepare and provide a series of resource materials to orient departments and agencies to assessment-related resources and expertise outside of the State:</p> <ul style="list-style-type: none"><li>a. "List of assessment-related resources" identifying the main text books, journals, and other printed materials for inclusion in a workplace library</li><li>b. "List of professional organizations and contacts" for professional networking and interaction</li><li>c. "List of available assessment-related training courses and college courses" for staff training and development</li><li>d. "List of outside recruitment sources" for the recruitment of trained Human Resources professionals to fill departmental examination unit vacancies</li></ul>	Spring 2003
<p>6. The SPB shall market existing services available on a reimbursable, staff-availability basis from the Test Validation and Construction (TV&amp;C) Unit:</p> <ul style="list-style-type: none"><li>a. "Shadow Analyst" service whereby departments and agencies contracting for expert TV&amp;C staff to conduct selection projects can utilize the presence of TV&amp;C staff to provide training and development opportunities for their examination analysts by assigning examination analysts to work closely with and observe TV&amp;C staff as project work is completed</li><li>b. "Rent-an-Analyst" service whereby departments and agencies can contract for expert TV&amp;C staff to (1) conduct complete or partial selection projects and/or (2) provide expert oversight and guidance to departmental analysts completing selection activities</li></ul>	Spring 2003

Recommendation	Proposed Implementation Timeline
7. The SPB shall design and offer a program to work closely with a select department or two to provide comprehensive guidance, support, training, and consultation in an effort to assist the department(s) in retooling the examination unit to meet program standards and expectations and become a model program for other departments/agencies to emulate	Spring 2003 <i>(program design)</i>  Spring/Summer 2003 <i>(program implementation)</i>
8. The SPB shall offer free training sessions to SPB staff, departmental examination managers and supervisors, and departmental examination analysts addressing the topics of job analysis, the development and use of structured interviews, and the use of job-related pass points	Spring/Summer 2003
9. The SPB shall pursue revisions to the classification specification for the Test Validation and Development Specialist series to ensure that the minimum qualifications established for each classification in the series are reflective of those qualifications minimally required to perform successfully in the job classifications	Fall 2003
10. Contingent upon the outcome of the Test Validation and Development Specialist classification series specification revision (as noted in recommendation #9), the SPB shall develop and administer a statewide, open examination for the Test Validation and Development Specialist I and II classifications	Late 2003 <i>(examination development)</i>  Early 2004 <i>(examination administration)</i>

Recommendation	Proposed Implementation Timeline
11. The SPB shall research the establishment of a specialized classification for the departmental examination manager/supervisor function(s) to provide for clear, job-related criteria upon which to classify and fill such positions as an option to ensuring the recruitment, selection, and retention of individuals with the required expertise	Late 2003/Early 2004
12. The SPB shall promulgate and expedite issuance of the <i>Merit Selection Manual: Policy and Practices</i> , which will address selection-related standards, expectations, guidelines, and best practices for the State's testing program	Beginning Fall 2002 and On-Going thereafter
13. Contingent upon the approval of these recommendations and program standards, the SPB shall conduct a review of the decentralized testing program following implementation of the program standards to determine the status of the program and departmental adherence to established standards and guidelines once the standards have been in effect for at least 24 months	Summer/Fall 2005

It should be emphasized that while more comprehensive measures to guide and oversee the decentralized testing program could be proposed, such additional measures would require additional staff and resource allocations for the SPB. Given the State's current budget realities, it is unlikely that the SPB would be allocated additional resources; thus, these proposed recommendations can be accomplished through the utilization of existing SPB staff and resources.

It is the goal of these recommendations to ensure that the State's Decentralized Testing Program is an effective, efficient means of assessing candidate job-related qualifications and to ensure that departmental testing processes and practices rely on the use of sound, job-related selection techniques.

## **Appendices**

**Survey Participants - Survey of State Government Testing Practices**

<b>Survey Participants</b>	
	Georgia
	Iowa
	Louisiana
	Maryland
	Michigan
	Minnesota
	New Jersey
	New York
	Ohio
	Pennsylvania
	Tennessee
	Wisconsin



***Decentralized Testing Questionnaire and Cover Memo***



February 15, 2002

Department  
Attn: Personnel Officer  
Address  
City, State Zip Code

### DECENTRALIZED TESTING QUESTIONNAIRE

The State Personnel Board (SPB) is conducting a comprehensive evaluation of the State's decentralized testing program. Since its introduction in 1981, the State's decentralized testing program has not been reviewed on a system-wide basis. In September 2001, the State Personnel Board passed a resolution calling for a system-wide review of the decentralized testing program. In accordance with that Board resolution, SPB staff have developed a comprehensive methodology by which to conduct the evaluation.

Encompassed in the project methodology is the dissemination of the attached questionnaire to all departments currently engaged in decentralized testing. The purpose of the questionnaire is to collect specific data relative to departmental testing practices, procedures, and results from those departments and agencies engaged in decentralized testing. The data collected will be aggregated and analyzed to evaluate the results of decentralized testing efforts within State service.

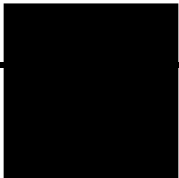
To ensure a thorough evaluation process, it is imperative that we collect adequate data from all State departments and agencies engaged in decentralized testing. To that end, we would appreciate your department's participation in our data collection efforts. Please note that we have selected March 8, 2002, as the return date for completed questionnaires.

Should you have questions regarding this questionnaire or the decentralized testing evaluation project, please feel free to contact Robin Bogdanich, (916) 654-6341, [rbogdanich@spb.ca.gov](mailto:rbogdanich@spb.ca.gov), or Shelley Langan, (916) 654-8538, [slangan@spb.ca.gov](mailto:slangan@spb.ca.gov), Policy Division staff.

Sincerely,

Mike Willihnganz  
Chief, Policy Division

Attachment



# DECENTRALIZED TESTING QUESTIONNAIRE

PROJECT CONDUCTED BY:  
POLICY DIVISION  
CALIFORNIA STATE PERSONNEL BOARD  
FEBRUARY 2002



For a copy of the questionnaire,  
please contact Robin Bogdanich at [rbogdanich@spb.ca.gov](mailto:rbogdanich@spb.ca.gov) or (916) 654-6341  
or Shelley Langan at [slangan@spb.ca.gov](mailto:slangan@spb.ca.gov) or (916) 654-8538.

**Thank you.**

**Departmental Interview Participants**

<b>Large Departments</b>	Department of Corrections
	Department of Forestry
	Department of Health Services
	Department of Motor Vehicles
	Department of Transportation
<b>Medium Departments</b>	California Public Employees' Retirement System
	Department of Industrial Relations
	Department of Parks and Recreation
	Department of Rehabilitation
	Department of Water Resources
<b>Small Departments</b>	Air Resources Board
	California State Personnel Board
	Department of Alcoholic Beverage Control
	Department of Fair Employment and Housing
	Prison Industry Authority
	Public Utilities Commission

**Departmental Interview Script**

---

**Decentralized Testing Evaluation Project**

February/March 2002

**Department Interview Notes**

---

Department

---

Department Representative/Title

---

Date of Interview

---

Interviewer

1. For each civil service examination process, who in your department decides what the *exam plan* (including the exam base, the scoring model for the process, and the components/instruments to be used) will be?
2. Does the examination staff in your department conduct job analyses?  
If not, why not?
3. Does the examination staff in your department work with subject matter experts (SMEs) while conducting examination planning, development, and administration activities?
  - a. If so:
    - Who are the SMEs (e.g., incumbents, first-level supervisors, etc.)?
    - Who in your department determines who the SMEs will be and how many of them will work with the examination staff in completing examination planning, development, and administration activities?
  - b. Typically, are adequate numbers of SMEs provided to your examination staff to participate in selection/examination activities?

4. What type of training or preparation related to conducting hiring interviews and making hiring decisions is provided to hiring supervisors in your department? (As applicable, who provides the training?)

5. What is currently *working* with the decentralized testing program?

6. What currently *is not working* with the decentralized testing program?

7. How could the decentralized testing program be *improved*?

How would such improvements affect the work activities of your department's examination staff?

8. What is your perception of the Board's role in the decentralized testing program?

9. What could the Board do, or do more of, to improve the decentralized testing program?

10. Are you currently sending your examination staff to the Selection Analyst Training Program classes offered through SPB's Technical Training Program?

If so:

a. What percentage of staff has attended one or more classes?

- If all staff not attending classes, why not?

b. Has your staff completed all of the classes in the Selection Analyst Training Program, and/or do you plan to have your staff complete all of the classes in the Program?

- If not, why not?

c. Are you and your staff applying the concepts and principles presented in the classes and curricula to your department's selection/examination activities?

- If so, how?
  - If not, why not?
-

## Union/Employee Advocacy Group Interview Participants

### Union Participants

Union	Bargaining Unit(s) Represented
American Federation of State, County and Municipal Employees (AFSCME)	Unit 19
California Association of Professional Scientists (CAPS)	Unit 10
California Association of Psychiatric Technicians (CAPT)	Unit 18
California Attorneys, Administrative Law Judges and Hearing Officers in State Employment (CASE)	Unit 2
California Department of Forestry Firefighters (CDF Firefighters)	Unit 8
California State Employees Association (CSEA)	Units 1, 3, 4, 11, 14, 15, 17, 20, and 21
California Union of Safety Employees (CAUSE)	Unit 7
International Union of Operating Engineers (IUOE)	Units 12 and 13

### Employee Advocacy Group Participants

- Association of California State Supervisors (ACSS)
- California Association of Managers and Supervisors (CAMS)



## Union/Employee Advocacy Group Interview Script

---

### Decentralized Testing Evaluation Project

February/March 2002

#### Union/Advocacy Group Interview Notes

---

Group

---

Group Representative/Title

---

Date of Interview

---

Interviewer

1. According to your membership, what is currently *working* with the State's decentralized testing program?
  
  2. According to your membership, what parameters or provisions of the State's decentralized testing program, if any, should be *discontinued*?
  
  3. What could the State Personnel Board do, or do more of, to improve the State's decentralized testing program?
-

***Departmental Selection Process Evaluation Criteria***

# Decentralized Testing Evaluation Project

March/April 2002

## Departmental Selection Process Evaluation Criteria

Examination Title	Department
Date of Administration	Reviewer Signature/Date of Review

General Information <i>[Check/fill-in as applicable.]</i>			
Examination Base	Open	<input type="checkbox"/>	Open, Non-Promotional
	Promotional	<input type="checkbox"/>	Open and Promotional
List Scoring Method Type	Nine-limited scores	<input type="checkbox"/>	Full range scoring
	Six-limited scores	<input type="checkbox"/>	"Three-limited" scores
Scoring Model	Instrument #1 (type and weighting) _____ Instrument #2 (type and weighting) _____ Instrument #3 (type and weighting) _____ Instrument #4 (type and weighting) _____		

Area of Review	Benchmarks/Evaluation Criteria <i>[Circle applicable rating.]</i>
<b>Job-relatedness</b>	
<b>Job analysis w/proper documentation to include:</b> <ul style="list-style-type: none"> <li>Essential tasks with importance and frequency ratings</li> <li>KSAs identified/rated as important to job success and required upon entry</li> <li>Linkage of essential tasks and important, required-upon-entry KSAs</li> <li>Documentation of SME participants</li> </ul>	<div style="display: flex; align-items: flex-start;"> <div style="width: 20px; text-align: center; font-weight: bold; margin-right: 10px;">3</div> <div>Job analysis includes all required documentation requirements, as outlined at left, and meets legal requirements. The task and/or KSA statements are thoroughly worded, and the wording may exceed minimal acceptable standards.</div> </div> <div style="display: flex; align-items: flex-start; margin-top: 10px;"> <div style="width: 20px; text-align: center; font-weight: bold; margin-right: 10px;">2</div> <div>Job analysis includes all required documentation requirements, as outlined at left. Job analysis meets legal requirements. The task and/or KSA statements may not be thoroughly worded, although the wording meets minimal acceptable standards.</div> </div> <div style="display: flex; align-items: flex-start; margin-top: 10px;"> <div style="width: 20px; text-align: center; font-weight: bold; margin-right: 10px;">1</div> <div>Job analysis fails to conform to legally required documentation requirements. Some or all of the documentation listed at left is missing from job analysis.</div> </div> <div style="display: flex; align-items: flex-start; margin-top: 10px;"> <div style="width: 20px; text-align: center; font-weight: bold; margin-right: 10px;">0</div> <div>Minimal or no attempt made to conduct professionally defined job analysis.</div> </div>

Area of Review	Benchmarks/Evaluation Criteria <i>[Circle applicable rating.]</i>
<b>Job-relatedness (continued)</b>	
<b>Job analytic data to include:</b> <ul style="list-style-type: none"> <li>• Tasks identified as essential or important to successful job performance</li> <li>• KSAs identified/rated as important to job success and required upon entry</li> <li>• Task/KSA linkage</li> <li>• Documentation of SME participants</li> </ul>	<p><b>3</b> Job analytic data includes all of the documentation requirements outlined at left. The data collection/analysis process goes beyond the class specification task/KSA wording in an effort to fully document the task/KSA requirements. The task and/or KSA statements may not be thoroughly worded, although the wording meets minimal acceptable standards (e.g., the tasks are complete behavioral statements, the KSAs are operationally defined).</p> <p><b>2</b> Job analytic data includes all of the documentation requirements outlined at left. The data collection/analysis process goes beyond the class specification task/KSA wording in an effort to fully document the task/KSA requirements. The wording of the tasks and/or KSAs may not be thorough nor meet minimal standards (e.g., the tasks may not be complete behavioral statements, or the KSAs may not be operationally defined).</p> <p><b>1</b> Job analytic data fails to include all of the documentation requirements outlined at left. OR, the data collection/analysis process may include only the tasks/KSAs from the class specification without any effort to fully analyze the job.</p> <p><b>0</b> No attempt made to collect/analyze job analytic data.</p>
<b>Subject Matter Expert participation in job analysis or job analytic data collection/analysis process</b>	<p><b>2</b> SMEs participating in the job analysis/job analytic data collection/analysis process include both supervisors and incumbents, all of whom are documented to possess the necessary expertise. OR, if only supervisors or incumbents have participated, then a full explanation of why only one group was used is provided, and that explanation could be sustained under content validity strategy requirements.</p> <p><b>1</b> SMEs participating in the job analysis/job analytic data collection/analysis process fail to include adequate SMEs documented to possess the necessary expertise. Both supervisors and incumbents may have participated, or just one group or the other; however, there is no clear rationale documented for the selection of those SME participants, either in terms of level (e.g., incumbent, supervisor) or expertise.</p> <p><b>0</b> No attempt made to select expert SMEs, OR no attempt made to include both incumbents or supervisors, OR the rationale for not including both incumbents and supervisors not fully documented.</p>

Area of Review	Benchmarks/Evaluation Criteria [Circle applicable rating.]
<b>Examination Planning</b>	
<p><b>Decision for examination base to include:</b></p> <ul style="list-style-type: none"> <li>• Documentation of rationale for exam base</li> <li>• Indication that exam base as chosen/developed meets the “best interests” of the State for merit-based selection</li> </ul>	<p><b>2</b> Decision for specific exam base includes the documentation/rationale requirements outlined at left. The exam base provides for an adequate candidate pool whose expected qualifications will meet the selection needs of this selection process.</p> <p><b>1</b> Decision/rationale for specific exam base not documented clearly to substantiate that exam base will provide for an adequate candidate pool to meet the selection needs of this selection process.</p> <p><b>0</b> No attempt or limited attempt made to document or substantiate exam base decision, OR exam base as documented/substantiated will not provide for an adequate candidate pool to meet the selection needs of this selection process.</p>
<p><b>Instrument selection for selection process to include:</b></p> <ul style="list-style-type: none"> <li>• Selection based on job analysis or job analytic data</li> <li>• Selection based on the appropriateness of the instrument(s) to assess those important KSAs required upon entry</li> <li>• Selection based on the appropriateness of the instrument(s) for administration with the anticipated candidate group, considering the size of the group, recruitment needs, and administration timeline requirements.</li> </ul>	<p><b>3</b> The selection instrument(s) chosen for this selection process is (are) the <i>most</i> appropriate selection instrument(s) based upon all of the criteria outlined at left. The selection instruments selected will provide for a job-related assessment of candidate qualifications based upon the entry requirements of the job.</p> <p><b>2</b> The selection instrument(s) chosen for this selection process is (are) appropriate selection instrument(s) based upon all of the criteria outlined at left, although the rationale for selection may not include an emphasis on using the most appropriate instrument(s). The selection instruments selected will provide for a job-related assessment of candidate qualifications based upon the entry requirements of the job.</p> <p><b>1</b> The selection instrument(s) chosen for this selection process is (are) not appropriate selection instrument(s) based upon all of the criteria outlined at left, although some rationale or documentation for the instrument selection may exist. The selection instrument(s) selected will not provide for a job-related assessment of candidate qualifications based upon the entry requirements of the job.</p> <p><b>0</b> The instrument(s) chosen will fail to provide for a job-related assessment of candidate qualifications based upon the entry requirements of the job, and/or is (are) inappropriate for the selection need.</p>

Area of Review	Benchmarks/Evaluation Criteria [Circle applicable rating.]
<b>Examination Planning (continued)</b>	
<p><b>Scoring model decision based upon:</b></p> <ul style="list-style-type: none"> <li>• Job analysis or job analytic data</li> <li>• Administration issues/considerations, including anticipated candidate group size, recruitment needs, and administration timeline requirements.</li> </ul>	<p><b>3</b> The scoring model for this selection process is based upon all of the criteria outlined at left. The scoring model will provide for an <i>extensive</i>, job-related assessment of candidate qualifications based upon the entry requirements of the job.</p> <p><b>2</b> The scoring model for this selection process is based upon all of the criteria outlined at left. The scoring model will provide for a job-related assessment of candidate qualifications based upon the entry requirements of the job, although it may not provide for the most thorough assessment of candidate qualifications or for the most job-related quantification of candidate competency.</p> <p><b>0</b> The scoring model for this selection process will not provide for a job-related assessment of candidate qualifications based upon the entry requirements of the job, AND/OR it will not provide for the job-related quantification of candidate competency.</p>
<p><b>Subject Matter Expert participation in examination planning activities</b></p>	<p><b>2</b> SMEs participating in the examination planning activities include both supervisors and incumbents, all of whom are documented to possess the necessary expertise. OR, if only supervisors or incumbents have participated, then a full explanation of why only one group was used is provided, and that explanation could be sustained under content validity strategy requirements.</p> <p><b>1</b> SMEs participating in the examination planning activities fail to include adequate SMEs documented to possess the necessary expertise. Both supervisors and incumbents may have participated, or just one group or the other; however, there is no clear rationale documented for the selection of those SME participants, either in terms of level (e.g., incumbent, supervisor) or expertise.</p> <p><b>0</b> No attempt made to select expert SMEs, OR no attempt made to include both incumbents or supervisors, OR the rationale for not including both incumbents and supervisors not fully documented.</p>

Area of Review	Benchmarks/Evaluation Criteria <i>[Circle applicable rating.]</i>
<b>Examination Development – Written Examination</b>	
<b>Development methodology to include:</b> <ul style="list-style-type: none"> <li>• Based upon job analysis or job analytic data</li> <li>• SME input/participation</li> <li>• Pretesting</li> </ul>	<p><b>3</b> The development methodology used for this written exam meets all of the criteria outlined at left and will provide for the development of a content-valid instrument.</p> <p><b>2</b> The development methodology used for this written exam meets the first two bulleted criteria outlined at left and will provide for the development of a content-valid instrument.</p> <p><b>1</b> The development methodology used for this written exam includes at least one of the first two bulleted criteria outlined at left; however, the methodology will not result in the development of a content-valid instrument.</p> <p><b>0</b> The development methodology used for this written exam fails to include at least one of the first two bulleted criteria outlined and/or will not result in development of a content-valid instrument.</p>
<b>Subject Matter Expert participation in examination development activities</b>	<p><b>2</b> SMEs participating in examination development activities include both supervisors and incumbents, all of whom are documented to possess the necessary expertise. OR, if only supervisors or incumbents have participated, then a full explanation of why only one group was used is provided, and that explanation could be sustained under content validity strategy requirements.</p> <p><b>1</b> SMEs participating in examination development activities fail to include adequate SMEs documented to possess the necessary expertise. Both supervisors and incumbents may have participated, or just one group or the other; however, there is no clear rationale documented for the selection of those SME participants, either in terms of level (e.g., incumbent, supervisor) or expertise.</p> <p><b>0</b> No attempt made to select expert SMEs, OR no attempt made to include both incumbents or supervisors, OR the rationale for not including both incumbents and supervisors not fully documented.</p>
<b>Content of instrument to include consideration of:</b> <ul style="list-style-type: none"> <li>• KSAs included for assessment based upon job analysis or job analytic data</li> <li>• KSAs included for assessment appropriate for assessment with a written exam</li> <li>• Adequate subtest/segment length (i.e., at least 25 items)</li> <li>• Subtest(s)/segment(s) comprised of homogenous items</li> <li>• As applicable, multiple subtests/segments appropriately weighted based upon job analysis or job analytic data</li> </ul>	<p><b>2</b> Instrument content conforms to all criteria outlined at left. Basis of instrument content is sound and meets the requirements of content validity.</p> <p><b>1</b> Instrument content fails to conform to some of the criteria outlined at left.</p> <p><b>0</b> Instrument content fails to conform to all of the criteria outlined at left.</p>

Area of Review	Benchmarks/Evaluation Criteria <i>[Circle applicable rating.]</i>
<b>Examination Development – Written Examination (continued)</b>	
<b>Item construction to include:</b> <ul style="list-style-type: none"> <li>• Well-phrased stems</li> <li>• Parallel alternatives to the stems</li> <li>• Parallel keys and distracters</li> <li>• Clear alternatives</li> <li>• Alternatives of similar length and complexity</li> </ul>	<p><b>3</b> Items conform to the criteria outlined at left. Items are soundly developed and will sufficiently assess candidate possession of the KSA(s) being assessed.</p> <p><b>2</b> Items conform to the first four criteria outlined at left. For the most part, items are soundly developed and will sufficiently assess candidate possession of the KSA(s) being assessed.</p> <p><b>1</b> Items fail to conform to at least the first four criteria outlined at left. Items are not indicative of soundly developed items and will not adequately assess candidate possession of the KSA(s) being assessed.</p> <p><b>0</b> Items are poorly developed, meeting none of the criteria outlined at left. Items will not assess candidate possession of the KSA(s) being assessed.</p>
<b>Preliminary pass point setting activities to include:</b> <ul style="list-style-type: none"> <li>• Collection of MAC data from SMEs</li> <li>• Use of acceptable, recognized MAC methodology</li> <li>• Participation of a representative group of SMEs</li> <li>• Participation of appropriate SMEs, including documentation/rationale for those SMEs utilized, whether inclusive of incumbents and supervisors or one or other group</li> </ul>	<p><b>3</b> Preliminary pass point setting activities include all of the considerations outlined at left, and the activities represent extensive efforts toward the collection of data upon which to base a defensible job-related pass point.</p> <p><b>2</b> Preliminary pass point setting activities include all of the considerations outlined at left and provide for the adequate collection of preliminary pass point setting data to serve as the basis for a defensible job-related pass point.</p> <p><b>1</b> Preliminary pass point setting activities fail to include all of the considerations outlined at left; however, while inadequate in establishing a defensible job-related pass point, the activities do encompass some elements of job-relatedness.</p> <p><b>0</b> No preliminary pass point setting activities conducted, or activities not representative of a recognized job-related methodology.</p>
<b>Pass point setting methodology/documentation to include:</b> <ul style="list-style-type: none"> <li>• Evidence of job-relatedness (e.g., MAC data, SME input)</li> <li>• Use of examination statistics</li> <li>• Consideration of exam validity</li> <li>• Consideration of consequence of error</li> </ul>	<p><b>3</b> Pass point(s) set meets all of the criteria outlined at left and is (are) a defensible job-related pass point(s).</p> <p><b>2</b> Pass point(s) set meets all of the criteria outlined at left and is (are) a defensible job-related pass point, although it (they) may not be the best pass point(s) based upon the data.</p> <p><b>1</b> Pass point(s) set fails to meet some of the criteria outlined and would be difficult to defend on the basis of job-relatedness.</p> <p><b>0</b> Pass point(s) set fails to meet all of the criteria outlined and is (are) not job-related and could not be defended.</p>



Area of Review	Benchmarks/Evaluation Criteria <i>[Circle applicable rating.]</i>
<b>Examination Development – Interview</b>	
<b>Development methodology to include:</b> <ul style="list-style-type: none"> <li>• Based upon job analysis or job analytic data</li> <li>• SME input/participation</li> </ul>	<p><b>2</b> The development methodology used for this interview meets the criteria outlined at left and will provide for the development of a content-valid instrument.</p> <p><b>1</b> The development methodology used for this interview includes at least one of the criteria outlined at left; however, the methodology will not result in the development of a content-valid instrument.</p> <p><b>0</b> The development methodology used for this interview fails to meet the development criteria outlined and/or will not result in development of a content-valid instrument.</p>
<b>Subject Matter Expert participation in interview development process</b>	<p><b>2</b> SMEs participating in interview development activities include both supervisors and incumbents, all of whom are documented to possess the necessary expertise. OR, if only supervisors or incumbents have participated, then a full explanation of why only one group was used is provided, and that explanation could be sustained under content validity strategy requirements.</p> <p><b>1</b> SMEs participating in interview development activities fail to include adequate SMEs documented to possess the necessary expertise. Both supervisors and incumbents may have participated, or just one group or the other; however, there is no clear rationale documented for the selection of those SME participants, either in terms of level (e.g., incumbent, supervisor) or expertise.</p> <p><b>0</b> No attempt made to select expert SMEs, OR no attempt made to include both incumbents or supervisors, OR the rationale for not including both incumbents and supervisors not fully documented.</p>
<b>Content of instrument to include consideration of:</b> <ul style="list-style-type: none"> <li>• KSAs included for assessment based upon job analysis or job analytic data</li> <li>• KSAs included for assessment appropriate for assessment with an interview</li> <li>• Adequate number of questions/dimensions (i.e., at least 7 questions/dimensions)</li> <li>• Appropriate question types used to assess those KSAs included for assessment in the interview</li> <li>• As applicable, specific differential weighting of questions appropriately based upon job analysis or job analytic data</li> </ul>	<p><b>3</b> Instrument content conforms to all criteria outlined at left. Basis of instrument content is sound, job-related, and meets the requirements of content validity. Content of the instrument may utilize elaborate question types or design parameters, distinguishing it as a complex interview which will better assess candidate qualifications than a minimally acceptable interview process.</p> <p><b>2</b> Instrument content conforms to all criteria outlined at left. Basis of instrument content is sound and meets the requirements of content validity.</p> <p><b>1</b> Instrument content fails to conform to some of the criteria outlined at left, and/or instrument fails to meet the requirements of content validity.</p> <p><b>0</b> Instrument content fails to conform to all of the criteria outlined at left.</p>

Area of Review	Benchmarks/Evaluation Criteria <i>[Circle applicable rating.]</i>
<b>Examination Development – Interview (continued)</b>	
<b>Question construction to include:</b> <ul style="list-style-type: none"> <li>• Well-phrased questions</li> <li>• Clearly worded questions that will elicit specific information from candidates</li> <li>• Questions with specific answers and/or responses</li> <li>• Question types that adequately measure the KSAs identified for assessment</li> </ul>	<p><b>3</b> Questions conform to the criteria outlined at left. Questions are soundly developed and will sufficiently assess candidate possession of the KSA(s) being assessed.</p> <p><b>2</b> Questions conform to the criteria outlined at left, although the types of questions may not represent the best means of assessing the KSAs at-hand (e.g., use of situational questions vs. behavioral questions, use of job knowledge questions vs. situational questions, etc.). For the most part, questions are soundly developed and will sufficiently assess candidate possession of the KSA(s) being assessed.</p> <p><b>1</b> Questions fail to conform to the criteria outlined at left. Questions are not soundly developed questions and/or will not adequately assess candidate possession of the KSA(s) being assessed.</p> <p><b>0</b> Questions are poorly developed, meeting none of the criteria outlined at left. Questions will not assess candidate possession of the KSA(s) being assessed.</p>
<b>Scoring criteria to include:</b> <ul style="list-style-type: none"> <li>• Structured scoring criteria that differentiates candidate performance</li> <li>• A clear distinction/determination of minimal acceptable competence</li> <li>• Job-related tie between scoring criteria/points and suggested responses/answers</li> <li>• Anchored rating scale with benchmark answers</li> </ul>	<p><b>3</b> Scoring criteria provides for a structured, standardized, job-related means of rating candidate responses. Scoring criteria includes all of the bulleted criteria outlined at left, meeting the standard for 'structured interviews.'</p> <p><b>2</b> Scoring criteria provides for a structured, standardized, job-related means of rating candidate responses. Scoring criteria includes all of the bulleted criteria outlined at left, <i>except</i> for the last bullet.</p> <p><b>1</b> Scoring criteria fails to provide for a structured, standardized, job-related means of rating candidate responses. Scoring criteria does not include at least the first three bulleted criteria.</p> <p><b>0</b> Limited or no scoring criteria developed.</p>
<b>Pass point setting methodology/documentation to include:</b> <ul style="list-style-type: none"> <li>• Adherence to pre-defined pass point if anchored rating scale/rating criteria used</li> <li>• Evidence of job-relatedness (e.g., SME input)</li> </ul>	<p><b>2</b> Pass point set meets the criteria outlined at left and is a defensible job-related pass point.</p> <p><b>1</b> Pass point set fails to meet the criteria outlined and would be difficult to defend on the basis of job-relatedness.</p> <p><b>0</b> Pass point set fails to meet the criteria outlined and is not a defensible job-related pass point.</p>

Area of Review	Benchmarks/Evaluation Criteria <i>[Circle applicable rating.]</i>
<b>Examination Development – Performance Test/Work Sample Exercise</b>	
<b>Development methodology to include:</b> <ul style="list-style-type: none"> <li>• Based upon job analysis or job analytic data</li> <li>• SME input/participation</li> <li>• Pretesting, as appropriate</li> </ul>	<p><b>3</b> The development methodology used for this performance-based test meets the criteria outlined at left and will provide for the development of a content-valid instrument.</p> <p><b>2</b> The development methodology used for this performance-based test includes the first two bulleted criteria outlined at left and will provide for the development of a content-valid instrument.</p> <p><b>1</b> The development methodology used for this performance-based test fails to include the first two bulleted criteria outlined and/or will not result in development of a content-valid instrument.</p> <p><b>0</b> The development methodology used for this performance-based test fails to include at least one of the first two bulleted criteria outlined and/or will not result in development of a content-valid instrument.</p>
<b>Subject Matter Expert participation in test development process</b>	<p><b>2</b> SMEs participating in test development activities include both supervisors and incumbents, all of whom are documented to possess the necessary expertise. OR, if only supervisors or incumbents have participated, then a full explanation of why only one group was used is provided, and that explanation could be sustained under content validity strategy requirements.</p> <p><b>1</b> SMEs participating in test development activities fail to include adequate SMEs documented to possess the necessary expertise. Both supervisors and incumbents may have participated, or just one group or the other; however, there is no clear rationale documented for the selection of those SME participants, either in terms of level (e.g., incumbent, supervisor) or expertise.</p> <p><b>0</b> No attempt made to select expert SMEs, OR no attempt made to include both incumbents or supervisors, OR the rationale for not including both incumbents and supervisors not fully documented.</p>

Area of Review	Benchmarks/Evaluation Criteria <i>[Circle applicable rating.]</i>
<b>Examination Development – Performance Test/Work Sample Exercise (continued)</b>	
<b>Content of instrument</b> <ul style="list-style-type: none"> <li>• KSAs included for assessment based upon job analysis or job analytic data</li> <li>• KSAs included for assessment are appropriate for assessment with a performance-based instrument</li> <li>• Assessment of an adequate sample of KSAs/behaviors</li> <li>• As applicable, specific differential weighting of questions appropriately based upon job analysis or job analytic data</li> </ul>	<p><b>3</b> Instrument content conforms to all criteria outlined at left. Basis of instrument content is sound, job-related, and meets the requirements of content validity. Content of the instrument may utilize elaborate exercises, distinguishing it as a complex performance test which will better assess candidate qualifications than a minimally acceptable interview process.</p> <p><b>2</b> Instrument content conforms to all criteria outlined at left. Basis of instrument content is sound and meets the requirements of content validity.</p> <p><b>1</b> Instrument content fails to conform to some of the criteria outlined at left, and/or instrument fails to meet the requirements of content validity.</p> <p><b>0</b> Instrument content fails to conform to all of the criteria outlined at left.</p>
<b>Instrument construction to include:</b> <ul style="list-style-type: none"> <li>• Exercise(s)/behaviors designed to require performance at the same level and complexity as required on the job</li> <li>• Well-developed exercises/tasks that measure specifically identified KSAs</li> <li>• Well-developed exercises/tasks that will elicit specific candidate behavior(s)/response</li> <li>• Inclusion of standardized candidate materials, tools, equipment, and instructions, as applicable</li> </ul>	<p><b>3</b> Exercise(s) conform(s) to the criteria outlined at left. Exercise(s) are soundly developed and will sufficiently assess candidate possession of the KSA(s) being assessed.</p> <p><b>2</b> Exercise(s) conform to the criteria outlined at left, although the type(s) of exercise(s) may not represent the best means of assessing the KSAs at-hand (e.g., the use of less-demanding exercises vs. more demanding behavioral-based exercises). For the most part, exercise(s) is (are) soundly developed and will sufficiently assess candidate possession of the KSA(s) being assessed.</p> <p><b>1</b> Exercise(s) fail(s) to conform to the criteria outlined at left. Exercise(s) are not soundly developed and/or will not adequately assess candidate possession of the KSA(s) being assessed.</p> <p><b>0</b> Questions are poorly developed, meeting none of the criteria outlined at left. Questions will not assess candidate possession of the KSA(s) being assessed.</p>

Area of Review	Benchmarks/Evaluation Criteria [Circle applicable rating.]
<b>Examination Development – Performance Test/Work Sample Exercise (continued)</b>	
<b>Scoring criteria to include:</b> <ul style="list-style-type: none"> <li>• Structured scoring criteria that differentiates candidate performance</li> <li>• A clear distinction/determination of minimal acceptable competence</li> <li>• Job-related tie between scoring criteria/points and suggested responses/answers</li> <li>• Anchored rating scale with benchmark answers</li> </ul>	<p><b>3</b> Scoring criteria provides for a structured, standardized, job-related means of rating candidate performance/behavior(s). Scoring criteria includes all of the bulleted criteria outlined at left.</p> <p><b>2</b> Scoring criteria provides for a structured, standardized, job-related means of rating candidate performance/behavior(s). Scoring criteria includes all of the bulleted criteria outlined at left, <i>except</i> for the last bullet.</p> <p><b>1</b> Scoring criteria fails to provide for a structured, standardized, job-related means of rating candidate performance/behavior(s). Scoring criteria does not include at least the first three bulleted criteria.</p> <p><b>0</b> Limited or no scoring criteria developed.</p>
<b>Preliminary pass point setting activities to include:</b> <ul style="list-style-type: none"> <li>• Collection of MAC data from SMEs</li> <li>• Use of acceptable, recognized MAC methodology</li> <li>• Participation of a representative group of SMEs</li> <li>• Participation of appropriate SMEs, including documentation/rationale for those SMEs utilized, whether inclusive of incumbents and supervisors or one or other group</li> </ul>	<p><b>3</b> Preliminary pass point setting activities include all of the considerations outlined at left, and the activities represent extensive efforts toward the collection of data upon which to base a defensible job-related pass point.</p> <p><b>2</b> Preliminary pass point setting activities include all of the considerations outlined at left and provide for the adequate collection of preliminary pass point setting data to serve as the basis for a defensible job-related pass point.</p> <p><b>1</b> Preliminary pass point setting activities fail to include all of the considerations outlined at left; however, while inadequate in establishing a defensible job-related pass point, the activities do encompass some elements of job-relatedness.</p> <p><b>0</b> No preliminary pass point setting activities conducted, or activities not representative of a recognized job-related methodology.</p>
<b>Pass point setting methodology/documentation to include:</b> <ul style="list-style-type: none"> <li>• Evidence of job-relatedness (e.g., MAC data, SME input)</li> <li>• Use of examination statistics</li> <li>• Consideration of exam validity</li> <li>• Consideration of consequence of error</li> <li>• Adherence to pre-defined pass point <i>if</i> anchored rating scale/rating criteria used</li> </ul>	<p><b>3</b> Pass point(s) set meets all of the criteria outlined at left and is (are) a defensible job-related pass point(s).</p> <p><b>2</b> Pass point(s) set meets all of the criteria outlined at left and is (are) a defensible job-related pass point, although it (they) may not be the best pass point(s) based upon the data.</p> <p><b>1</b> Pass point(s) set fails to meet some of the criteria outlined and would be difficult to defend on the basis of job-relatedness.</p> <p><b>0</b> Pass point(s) set fails to meet all of the criteria outlined and is (are) not job-related and could not be defended.</p>

Area of Review	Benchmarks/Evaluation Criteria <i>[Circle applicable rating.]</i>
<b>Examination Development – Supplemental Application</b>  <input type="checkbox"/> Behavioral Consistency Method SAF  <input type="checkbox"/> Self-Rating Point Method SAF	
<b>Development methodology to include:</b> <ul style="list-style-type: none"> <li>• Based upon job analysis or job analytic data</li> <li>• SME input/participation</li> </ul>	<p><b>2</b> The development methodology used for this SAF meets the criteria outlined at left and will provide for the development of a content-valid instrument.</p> <p><b>1</b> The development methodology used for this SAF includes one of the two bulleted criteria; however, the methodology will not provide for the development of a content-valid instrument.</p> <p><b>0</b> The development methodology used for this SAF fails to meet the development criteria outlined and/or will not result in development of a content-valid instrument.</p>
<b>Subject Matter Expert participation in SAF development process</b>	<p><b>2</b> SMEs participating in SAF development activities include both supervisors and incumbents, all of whom are documented to possess the necessary expertise. OR, if only supervisors or incumbents have participated, then a full explanation of why only one group was used is provided, and that explanation could be sustained under content validity strategy requirements.</p> <p><b>1</b> SMEs participating in SAF development activities fail to include adequate SMEs documented to possess the necessary expertise. Both supervisors and incumbents may have participated, or just one group or the other; however, there is no clear rationale documented for the selection of those SME participants, either in terms of level (e.g., incumbent, supervisor) or expertise.</p> <p><b>0</b> No attempt made to select expert SMEs, OR no attempt made to include both incumbents or supervisors, OR the rationale for not including both incumbents and supervisors not fully documented.</p>

Area of Review	Benchmarks/Evaluation Criteria <i>[Circle applicable rating.]</i>
<b>Examination Development – Supplemental Application (continued)</b>	
<b>Content of instrument</b> <ul style="list-style-type: none"> <li>• KSAs included for assessment based upon job analysis or job analytic data</li> <li>• KSAs included for assessment are appropriate for assessment with an SAF</li> <li>• Assessment of an adequate sample of KSAs/behaviors</li> <li>• As applicable, specific differential weighting of questions/items appropriately based upon job analysis or job analytic data</li> </ul>	<p><b>3</b> Instrument content conforms to all criteria outlined at left. Basis of instrument content is sound, job-related, and meets the requirements of content validity. Content of the instrument may utilize elaborate items or means of accessing candidate qualifications, distinguishing it as a complex SAF which will better assess candidate qualifications than a minimally acceptable SAF.</p> <p><b>2</b> Instrument content conforms to all criteria outlined at left. Basis of instrument content is sound and meets the requirements of content validity.</p> <p><b>1</b> Instrument content fails to conform to some of the criteria outlined at left, and/or instrument fails to meet the requirements of content validity.</p> <p><b>0</b> Instrument content fails to conform to all of the criteria outlined at left.</p>
<b>Question/item construction to include:</b> <ul style="list-style-type: none"> <li>• Well-phrased questions</li> <li>• Clearly worded questions that will elicit specific information from candidates</li> <li>• Questions with specific answers and/or responses</li> <li>• Question types that adequately measure the KSAs identified for assessment</li> </ul>	<p><b>3</b> Questions/items conform to the criteria outlined at left. Questions/items are soundly developed and will sufficiently assess candidate possession of the KSA(s) being assessed.</p> <p><b>2</b> Questions conform to the criteria outlined at left, although the types of questions may not represent the best means of assessing the KSAs at-hand (e.g., use of self-rating questions vs. behavioral questions, etc.). For the most part, questions/items are soundly developed and will sufficiently assess candidate possession of the KSA(s) being assessed.</p> <p><b>1</b> Questions/items fail to conform to the criteria outlined at left. Questions/items are not soundly developed and/or will not adequately assess candidate possession of the KSA(s) being assessed.</p> <p><b>0</b> Questions/items are poorly developed, meeting none of the criteria outlined at left, and will not assess candidate possession of the KSA(s) being assessed.</p>

Area of Review	Benchmarks/Evaluation Criteria [Circle applicable rating.]
<b>Examination Development – Supplemental Application (continued)</b>	
<b>Scoring criteria to include:</b> <ul style="list-style-type: none"> <li>• Structured scoring criteria that differentiates candidate performance</li> <li>• A clear distinction/determination of minimal acceptable competence</li> <li>• Job-related tie between scoring criteria/points and suggested responses/answers</li> <li>• Anchored rating scale with benchmark answers</li> </ul>	<p><b>3</b> Scoring criteria provides for a structured, standardized, job-related means of rating candidate experience/qualification(s). Scoring criteria includes all of the bulleted criteria outlined at left.</p> <p><b>2</b> Scoring criteria provides for a structured, standardized, job-related means of rating candidate experience/qualification(s). Scoring criteria includes all of the bulleted criteria outlined at left, <i>except</i> for the last bullet.</p> <p><b>1</b> Scoring criteria fails to provide for a structured, standardized, job-related means of rating candidate experience/qualifications. Scoring criteria does not include at least the first three bulleted criteria.</p> <p><b>0</b> Limited or no scoring criteria developed.</p>
<b>Pass point setting methodology/documentation to include:</b> <ul style="list-style-type: none"> <li>• Evidence of job-relatedness (e.g., MAC data, SME input)</li> <li>• Use of examination statistics</li> <li>• Consideration of exam validity</li> <li>• Consideration of consequence of error</li> <li>• Adherence to pre-defined pass point <i>if</i> anchored rating scale/rating criteria used</li> </ul>	<p><b>3</b> Pass point set meets all of the criteria outlined at left and is a defensible job-related pass point.</p> <p><b>2</b> Pass point set meets all of the criteria outlined at left and is a defensible job-related pass point, although it may not be the best pass point based upon the data.</p> <p><b>1</b> Pass point set fails to meet some of the criteria outlined and would be difficult to defend on the basis of job-relatedness.</p> <p><b>0</b> Pass point set fails to meet all of the criteria outlined and is not job-related and could not be defended.</p>



Area of Review	Benchmarks/Evaluation Criteria [Circle applicable rating.]
<b>Examination Administration – Instrument #1:</b> _____	
<b>Candidate Instructions</b> <ul style="list-style-type: none"> <li>Clearly worded to outline candidate performance requirements</li> <li>Standardized for all candidates</li> <li>Applicable/related to <i>expectations</i> of candidate performance based upon scoring criteria</li> </ul>	<p><b>3</b> Candidate instructions, meeting all of the criteria outlined at left, are thoroughly worded, well-developed, and clearly presented to the candidates.</p> <p><b>2</b> Adequate candidate instructions meeting all of the criteria outlined at left are presented to the candidates, although the instructions could be more complete/more thorough.</p> <p><b>1</b> Candidate instructions are not well-developed or are not clearly presented to the candidates. Instructions fail to meet the criteria outlined at left.</p> <p><b>0</b> Limited or no instructions provided to candidates.</p>
<b>Rater training/preparation to include, as applicable:</b> <ul style="list-style-type: none"> <li>Explanation of expectation of candidate performance</li> <li>Explanation of testing instrument</li> <li>Explanation of scoring criteria and how to utilize/apply criteria</li> <li>Ample preparation time for raters to review materials and participate in rater training</li> <li>Review of expectations of rater performance/behavior</li> </ul>	<p><b>3</b> Extensive rater training and preparation provided to raters to include all of the criteria outlined at left.</p> <p><b>2</b> Sufficient rater training and preparation provided to raters to include all of the criteria outlined at left.</p> <p><b>1</b> Insufficient rater training and/or preparation provided to raters; such training does not include all of the criteria outlined at left.</p> <p><b>0</b> Limited or no rater training and/or preparation provided to raters.</p>
<b>Rater qualifications and appropriateness of participation</b>	<p><b>2</b> SMEs participating in rating/scoring activities are all documented to possess the necessary expertise to appropriately rate candidate performance. A full explanation of why/how the SMEs were selected is provided. The SMEs may include both supervisors and incumbents.</p> <p><b>1</b> SMEs participating in rating/scoring activities may not possess the necessary expertise to appropriately rate candidate performance, or insufficient documentation is provided to adequately establish the expertise of the SMEs. OR, an explanation of why/how the SMEs were selected may not be provided, or that explanation fails to adequately document the selection rationale. The SMEs may include both supervisors and incumbents.</p> <p><b>0</b> No attempt made to select expert SMEs, OR no attempt made to document the rationale for selecting the SME participants.</p>

Area of Review	Benchmarks/Evaluation Criteria <i>[Circle applicable rating.]</i>	
<b>Examination Administration – Instrument #1 (continued)</b>		
<b>Application of rating/scoring criteria when instrument administered/scored</b>	<b>2</b>       <b>1</b>       <b>0</b>	Scoring criteria is consistently applied as developed in rating candidate performance.  Scoring criteria is not consistently applied as developed in rating candidate performance, although the criteria is used to rate performance for some of the candidates.  Scoring criteria is not applied as developed in rating candidate performance.
<b>Examination Statistics – Instrument #1: _____</b>		
<b>Quality of available examination statistics to include:</b> <ul style="list-style-type: none"> <li>• Item analysis, as applicable <ul style="list-style-type: none"> <li>• Run by exam segment/subtest</li> </ul> </li> <li>• Raw score tab. report(s) <ul style="list-style-type: none"> <li>• Run by exam segment/subtest if separate segment/subtest pass points set (or to be set)</li> </ul> </li> <li>• Examination control documentation</li> <li>• Bottom line hiring reports indicative of final examination results</li> </ul>	<b>2</b>       <b>1</b>       <b>0</b>	Adequate and correct exam statistics/reports run, available, and utilized appropriately.  Adequate and/or correct exam statistics/reports either not run, available, and/or utilized appropriately.  Insufficient record(s) or no record(s) of exam statistics/reports.  <b>Comments:</b>
<b>Use of statistics/interpretation of statistics in finalizing examination scoring, determining pass points, processing scores, etc., to include:</b> <ul style="list-style-type: none"> <li>• Item analysis used to identify and rekey miskeyed items</li> <li>• Item analysis used to identify and delete poor-performing items</li> <li>• Raw score tab. used in pass point setting activities.</li> </ul>	<b>2</b>       <b>1</b>       <b>0</b>	Examination statistics and reports used appropriately and correctly in making examination scoring decisions, including determining pass points, and interpreting item analysis results.  Examination statistics and reports used inappropriately and/or incorrectly in making examination scoring decisions, determining pass points, and/or interpreting item analysis results.  Limited access/use of examination statistics and reports and/or no access/use of examination statistics and reports in making examination scoring decisions.
<b>Is there evidence of adverse impact with this instrument's use?</b>	<b>Yes</b>       <b>No</b>	<b>Comments:</b>
<b>Can such evidence of adverse impact be defended on the basis of business necessity?</b>	<b>N/A</b>       <b>Yes</b>       <b>No</b>	<b>Comments:</b>

Area of Review	Benchmarks/Evaluation Criteria [Circle applicable rating.]
<b>Examination Administration – Instrument #2:</b> _____	
<b>Candidate Instructions</b> <ul style="list-style-type: none"> <li>Clearly worded to outline candidate performance requirements</li> <li>Standardized for all candidates</li> <li>Applicable/related to <i>expectations</i> of candidate performance based upon scoring criteria</li> </ul>	<p><b>3</b> Candidate instructions, meeting all of the criteria outlined at left, are thoroughly worded, well-developed, and clearly presented to the candidates.</p> <p><b>2</b> Adequate candidate instructions meeting all of the criteria outlined at left are presented to the candidates, although the instructions could be more complete/more thorough.</p> <p><b>1</b> Candidate instructions are not well-developed or are not clearly presented to the candidates. Instructions fail to meet the criteria outlined at left.</p> <p><b>0</b> Limited or no instructions provided to candidates.</p>
<b>Rater training/preparation to include, as applicable:</b> <ul style="list-style-type: none"> <li>Explanation of expectation of candidate performance</li> <li>Explanation of testing instrument</li> <li>Explanation of scoring criteria and how to utilize/apply criteria</li> <li>Ample preparation time for raters to review materials and participate in rater training</li> <li>Review of expectations of rater performance/behavior</li> </ul>	<p><b>3</b> Extensive rater training and preparation provided to raters to include all of the criteria outlined at left.</p> <p><b>2</b> Sufficient rater training and preparation provided to raters to include all of the criteria outlined at left.</p> <p><b>1</b> Insufficient rater training and/or preparation provided to raters; such training does not include all of the criteria outlined at left.</p> <p><b>0</b> Limited or no rater training and/or preparation provided to raters.</p>
<b>Rater qualifications and appropriateness of participation</b>	<p><b>2</b> SMEs participating in rating/scoring activities are all documented to possess the necessary expertise to appropriately rate candidate performance. A full explanation of why/how the SMEs were selected is provided. The SMEs may include both supervisors and incumbents.</p> <p><b>1</b> SMEs participating in rating/scoring activities may not possess the necessary expertise to appropriately rate candidate performance, or insufficient documentation is provided to adequately establish the expertise of the SMEs. OR, an explanation of why/how the SMEs were selected may not be provided, or that explanation fails to adequately document the selection rationale. The SMEs may include both supervisors and incumbents.</p> <p><b>0</b> No attempt made to select expert SMEs, OR no attempt made to document the rationale for selecting the SME participants.</p>

Area of Review	Benchmarks/Evaluation Criteria <i>[Circle applicable rating.]</i>	
<b>Examination Administration – Instrument #2 (continued)</b>		
<b>Application of rating/scoring criteria when instrument administered/scored</b>	<b>2</b>       <b>1</b>       <b>0</b>	Scoring criteria is consistently applied as developed in rating candidate performance.  Scoring criteria is not consistently applied as developed in rating candidate performance, although the criteria is used to rate performance for some of the candidates.  Scoring criteria is not applied as developed in rating candidate performance.
<b>Examination Statistics – Instrument #2: _____</b>		
<b>Quality of available examination statistics to include:</b> <ul style="list-style-type: none"> <li>• Item analysis, as applicable               <ul style="list-style-type: none"> <li>• Run by exam segment/subtest</li> </ul> </li> <li>• Raw score tab. report(s)               <ul style="list-style-type: none"> <li>• Run by exam segment/subtest if separate segment/subtest pass points set (or to be set)</li> </ul> </li> <li>• Examination control documentation</li> <li>• Bottom line hiring reports indicative of final examination results</li> </ul>	<b>2</b>       <b>1</b>       <b>0</b>	Adequate and correct exam statistics/reports run, available, and utilized appropriately.  Adequate and/or correct exam statistics/reports either not run, available, and/or utilized appropriately.  Insufficient record(s) or no record(s) of exam statistics/reports.  <b>Comments:</b>
<b>Use of statistics/interpretation of statistics in finalizing examination scoring, determining pass points, processing scores, etc., to include:</b> <ul style="list-style-type: none"> <li>• Item analysis used to identify and rekey miskeyed items</li> <li>• Item analysis used to identify and delete poor-performing items</li> <li>• Raw score tab. used in pass point setting activities.</li> </ul>	<b>2</b>       <b>1</b>       <b>0</b>	Examination statistics and reports used appropriately and correctly in making examination scoring decisions, including determining pass points, and interpreting item analysis results.  Examination statistics and reports used inappropriately and/or incorrectly in making examination scoring decisions, determining pass points, and/or interpreting item analysis results.  Limited access/use of examination statistics and reports and/or no access/use of examination statistics and reports in making examination scoring decisions.
<b>Is there evidence of adverse impact with this instrument's use?</b>	<b>Yes</b>       <b>No</b>	<b>Comments:</b>
<b>Can such evidence of adverse impact be defended on the basis of business necessity?</b>	<b>N/A</b>       <b>Yes</b>       <b>No</b>	<b>Comments:</b>

Area of Review	Benchmarks/Evaluation Criteria [Circle applicable rating.]
<b>Examination Administration – Instrument #3:</b> _____	
<b>Candidate Instructions</b> <ul style="list-style-type: none"> <li>Clearly worded to outline candidate performance requirements</li> <li>Standardized for all candidates</li> <li>Applicable/related to <i>expectations</i> of candidate performance based upon scoring criteria</li> </ul>	<p><b>3</b> Candidate instructions, meeting all of the criteria outlined at left, are thoroughly worded, well-developed, and clearly presented to the candidates.</p> <p><b>2</b> Adequate candidate instructions meeting all of the criteria outlined at left are presented to the candidates, although the instructions could be more complete/more thorough.</p> <p><b>1</b> Candidate instructions are not well-developed or are not clearly presented to the candidates. Instructions fail to meet the criteria outlined at left.</p> <p><b>0</b> Limited or no instructions provided to candidates.</p>
<b>Rater training/preparation to include, as applicable:</b> <ul style="list-style-type: none"> <li>Explanation of expectation of candidate performance</li> <li>Explanation of testing instrument</li> <li>Explanation of scoring criteria and how to utilize/apply criteria</li> <li>Ample preparation time for raters to review materials and participate in rater training</li> <li>Review of expectations of rater performance/behavior</li> </ul>	<p><b>3</b> Extensive rater training and preparation provided to raters to include all of the criteria outlined at left.</p> <p><b>2</b> Sufficient rater training and preparation provided to raters to include all of the criteria outlined at left.</p> <p><b>1</b> Insufficient rater training and/or preparation provided to raters; such training does not include all of the criteria outlined at left.</p> <p><b>0</b> Limited or no rater training and/or preparation provided to raters.</p>
<b>Rater qualifications and appropriateness of participation</b>	<p><b>2</b> SMEs participating in rating/scoring activities are all documented to possess the necessary expertise to appropriately rate candidate performance. A full explanation of why/how the SMEs were selected is provided. The SMEs may include both supervisors and incumbents.</p> <p><b>1</b> SMEs participating in rating/scoring activities may not possess the necessary expertise to appropriately rate candidate performance, or insufficient documentation is provided to adequately establish the expertise of the SMEs. OR, an explanation of why/how the SMEs were selected may not be provided, or that explanation fails to adequately document the selection rationale. The SMEs may include both supervisors and incumbents.</p> <p><b>0</b> No attempt made to select expert SMEs, OR no attempt made to document the rationale for selecting the SME participants.</p>

Area of Review	Benchmarks/Evaluation Criteria <i>[Circle applicable rating.]</i>	
<b>Examination Administration – Instrument #3 (continued)</b>		
<b>Application of rating/scoring criteria when instrument administered/scored</b>	<b>2</b>       <b>1</b>       <b>0</b>	Scoring criteria is consistently applied as developed in rating candidate performance.  Scoring criteria is not consistently applied as developed in rating candidate performance, although the criteria is used to rate performance for some of the candidates.  Scoring criteria is not applied as developed in rating candidate performance.
<b>Examination Statistics – Instrument #3: _____</b>		
<b>Quality of available examination statistics to include:</b> <ul style="list-style-type: none"> <li>• Item analysis, as applicable               <ul style="list-style-type: none"> <li>• Run by exam segment/subtest</li> </ul> </li> <li>• Raw score tab. report(s)               <ul style="list-style-type: none"> <li>• Run by exam segment/subtest if separate segment/subtest pass points set (or to be set)</li> </ul> </li> <li>• Examination control documentation</li> <li>• Bottom line hiring reports indicative of final examination results</li> </ul>	<b>2</b>       <b>1</b>       <b>0</b>	Adequate and correct exam statistics/reports run, available, and utilized appropriately.  Adequate and/or correct exam statistics/reports either not run, available, and/or utilized appropriately.  Insufficient record(s) or no record(s) of exam statistics/reports.  <b>Comments:</b>
<b>Use of statistics/interpretation of statistics in finalizing examination scoring, determining pass points, processing scores, etc., to include:</b> <ul style="list-style-type: none"> <li>• Item analysis used to identify and rekey miskeyed items</li> <li>• Item analysis used to identify and delete poor-performing items</li> <li>• Raw score tab. used in pass point setting activities.</li> </ul>	<b>2</b>       <b>1</b>       <b>0</b>	Examination statistics and reports used appropriately and correctly in making examination scoring decisions, including determining pass points, and interpreting item analysis results.  Examination statistics and reports used inappropriately and/or incorrectly in making examination scoring decisions, determining pass points, and/or interpreting item analysis results.  Limited access/use of examination statistics and reports and/or no access/use of examination statistics and reports in making examination scoring decisions.
<b>Is there evidence of adverse impact with this instrument's use?</b>	<b>Yes</b>       <b>No</b>	<b>Comments:</b>
<b>Can such evidence of adverse impact be defended on the basis of business necessity?</b>	<b>N/A</b>       <b>Yes</b>       <b>No</b>	<b>Comments:</b>

Area of Review	Benchmarks/Evaluation Criteria [Circle applicable rating.]
<b>Examination Administration – Instrument #4:</b> _____	
<b>Candidate Instructions</b> <ul style="list-style-type: none"> <li>Clearly worded to outline candidate performance requirements</li> <li>Standardized for all candidates</li> <li>Applicable/related to <i>expectations</i> of candidate performance based upon scoring criteria</li> </ul>	<p><b>3</b> Candidate instructions, meeting all of the criteria outlined at left, are thoroughly worded, well-developed, and clearly presented to the candidates.</p> <p><b>2</b> Adequate candidate instructions meeting all of the criteria outlined at left are presented to the candidates, although the instructions could be more complete/more thorough.</p> <p><b>1</b> Candidate instructions are not well-developed or are not clearly presented to the candidates. Instructions fail to meet the criteria outlined at left.</p> <p><b>0</b> Limited or no instructions provided to candidates.</p>
<b>Rater training/preparation to include, as applicable:</b> <ul style="list-style-type: none"> <li>Explanation of expectation of candidate performance</li> <li>Explanation of testing instrument</li> <li>Explanation of scoring criteria and how to utilize/apply criteria</li> <li>Ample preparation time for raters to review materials and participate in rater training</li> <li>Review of expectations of rater performance/behavior</li> </ul>	<p><b>3</b> Extensive rater training and preparation provided to raters to include all of the criteria outlined at left.</p> <p><b>2</b> Sufficient rater training and preparation provided to raters to include all of the criteria outlined at left.</p> <p><b>1</b> Insufficient rater training and/or preparation provided to raters; such training does not include all of the criteria outlined at left.</p> <p><b>0</b> Limited or no rater training and/or preparation provided to raters.</p>
<b>Rater qualifications and appropriateness of participation</b>	<p><b>2</b> SMEs participating in rating/scoring activities are all documented to possess the necessary expertise to appropriately rate candidate performance. A full explanation of why/how the SMEs were selected is provided. The SMEs may include both supervisors and incumbents.</p> <p><b>1</b> SMEs participating in rating/scoring activities may not possess the necessary expertise to appropriately rate candidate performance, or insufficient documentation is provided to adequately establish the expertise of the SMEs. OR, an explanation of why/how the SMEs were selected may not be provided, or that explanation fails to adequately document the selection rationale. The SMEs may include both supervisors and incumbents.</p> <p><b>0</b> No attempt made to select expert SMEs, OR no attempt made to document the rationale for selecting the SME participants.</p>

Area of Review	Benchmarks/Evaluation Criteria <i>[Circle applicable rating.]</i>	
<b>Examination Administration – Instrument #4 (continued)</b>		
<b>Application of rating/scoring criteria when instrument administered/scored</b>	<b>2</b> Scoring criteria is consistently applied as developed in rating candidate performance.  <b>1</b> Scoring criteria is not consistently applied as developed in rating candidate performance, although the criteria is used to rate performance for some of the candidates.  <b>0</b> Scoring criteria is not applied as developed in rating candidate performance.	
<b>Examination Statistics – Instrument #4:</b> _____		
<b>Quality of available examination statistics to include:</b> <ul style="list-style-type: none"> <li>• Item analysis, as applicable <ul style="list-style-type: none"> <li>• Run by exam segment/subtest</li> </ul> </li> <li>• Raw score tab. report(s) <ul style="list-style-type: none"> <li>• Run by exam segment/subtest if separate segment/subtest pass points set (or to be set)</li> </ul> </li> <li>• Examination control documentation</li> <li>• Bottom line hiring reports indicative of final examination results</li> </ul>	<b>2</b> Adequate and correct exam statistics/reports run, available, and utilized appropriately.  <b>1</b> Adequate and/or correct exam statistics/reports either not run, available, and/or utilized appropriately.  <b>0</b> Insufficient record(s) or no record(s) of exam statistics/reports.  <b>Comments:</b>	
<b>Use of statistics/interpretation of statistics in finalizing examination scoring, determining pass points, processing scores, etc., to include:</b> <ul style="list-style-type: none"> <li>• Item analysis used to identify and rekey miskeyed items</li> <li>• Item analysis used to identify and delete poor-performing items</li> <li>• Raw score tab. used in pass point setting activities.</li> </ul>	<b>2</b> Examination statistics and reports used appropriately and correctly in making examination scoring decisions, including determining pass points, and interpreting item analysis results.  <b>1</b> Examination statistics and reports used inappropriately and/or incorrectly in making examination scoring decisions, determining pass points, and/or interpreting item analysis results.  <b>0</b> Limited access/use of examination statistics and reports and/or no access/use of examination statistics and reports in making examination scoring decisions.	
<b>Is there evidence of adverse impact with this instrument's use?</b>	<b>Yes</b> <i>Comments:</i>  <b>No</b>	
<b>Can such evidence of adverse impact be defended on the basis of business necessity?</b>	<b>N/A</b> <i>Comments:</i>  <b>Yes</b>  <b>No</b>	



Area of Review	Benchmarks/Evaluation Criteria [Circle applicable rating.]	
Final Scoring/Eligible List Creation		
Scoring Model – was scoring model adhered to in the computation of final candidate scores?	<div>2</div> <div>0</div>	<div>Scoring model was correctly and appropriately adhered to in the computation of final candidate scores and eligible list results.</div> <div>Scoring was incorrectly and/or inappropriately adhered to in the computation of final candidate scores and/or eligible list results.</div> <div>Comments:</div>
Where multiple weighted components comprised the selection process, were the scores from each component standardized and weighted accordingly to derive final candidate scores?	<div>2</div> <div>0</div>	<div>Component scores were correctly and appropriately standardized and weighted in computing final candidate scores and/or eligible list results.</div> <div>Component scores were not correctly/appropriately standardized and/or weighted in computing final candidate scores and/or eligible list results.</div> <div>Comments:</div>
Is there evidence of <i>bottom line</i> adverse impact?	<div>Yes</div> <div>No</div>	<div>Comments:</div>
Can such evidence of <i>bottom line</i> adverse impact be defended on the basis of business necessity at each step in the selection process to include: <ul style="list-style-type: none"><li>• Job-related rationale for selection process/exam plan</li><li>• Job-related rationale for instrument selection</li><li>• Job-related rationale for scoring model</li><li>• Job-related instrument content</li><li>• Job-related scoring criteria for instrument(s)</li><li>• Job-related pass point(s)</li></ul>	<div>N/A</div> <div>Yes</div> <div>No</div>	<div>Comments:</div>

**Comments/Additional Notes:**

---

---

---

---

---

---

---

---

---

---

**Scale:**

- 3 Exceeds minimal acceptable standard(s)
- 2 Sufficiently meets minimal acceptable standard(s)
- 1 Attempts, but fails to meet minimal acceptable standard(s)
- 0 Fails to meet minimal acceptable standard(s)

## **Groups Participating in the July 2002 State Personnel Board Hearing**

---

### **Groups Providing Oral Testimony during Hearing**

---

Association of California State Supervisors  
California Association of Psychiatric Technicians  
California Department of Forestry Firefighters Association  
California State Employees' Association  
County of Sacramento  
Department of Corrections  
Department of Health Services  
Examination Supervisors' Forum  
Small Personnel Information Network (SPIN)

---

---

### **Groups Providing Written Comments in conjunction with Hearing**

---

Ad Hoc Committee on Personnel  
Association of California State Supervisors  
California State Employees' Association  
Department of Justice  
Department of Mental Health  
Franchise Tax Board

---

**Hearing Notice for July 2002 Hearing before the State Personnel Board**



## CALIFORNIA STATE PERSONNEL BOARD

GRAY DAVIS, Governor

801 Capitol Mall • Sacramento, California 95814 • [www.spb.ca.gov](http://www.spb.ca.gov)



DATE: June 18, 2002

TO: ALL STATE AGENCIES AND EMPLOYEE ORGANIZATIONS

SUBJECT: HEARING ON THE STATUS OF THE STATE'S DECENTRALIZED TESTING PROGRAM

In compliance with a resolution adopted September 7, 2001, by the five-member State Personnel Board (the Board), entitled, *To Promote Sound Testing Practices*, the State Personnel Board (SPB) will hold a public hearing on July 9, 2002, to consider the staff's report on *The Status of the State's Decentralized Testing Program: Preliminary Findings and Recommendations*. The hearing will be held at 3:15 p.m., in Room 150, at the State Personnel Board, 801 Capitol Mall, Sacramento, California 95814. Those interested in providing oral testimony regarding the preliminary findings and recommendations outlined in the staff's report should notify Shelley Langan, Policy Division, at (916) 654-8538, by July 8, 2002. Those wishing to provide written testimony may submit their comments to Ms. Langan at the above address; written testimony should be received by close-of-business on July 8, 2002.

The Board resolution called for a comprehensive review and evaluation of the State's decentralized testing program, as well as a public hearing to solicit comments and input from program stakeholders. SPB staff has completed a comprehensive review of the program and issued preliminary findings and recommendations for comment at the above-noted public hearing. Preliminary findings to-date concluded that the State's decentralized testing program could be improved. A series of program improvement recommendations has been issued by SPB staff. Attached is an Executive Summary of the report of the preliminary findings and recommendations. Interested parties may access and download the complete report on SPB's Web site at [www.spb.ca.gov](http://www.spb.ca.gov). A limited number of printed copies of the report are available from the SPB's Policy Division. To request a copy, call (916) 654-8538. In addition, a limited number of copies will be available at the hearing. Following the hearing, SPB staff will prepare a final report and recommendations for consideration by the Board. The Board will then issue final recommendations relative to the Decentralized Testing Program.

Questions regarding this matter may be directed to Shelley Langan, Manager, Special Projects, Policy Division, at (916) 654-8538.

Walter Vaughn  
Executive Officer

Attachment

### Scoring Models and Instruments Comprising the Decentralized Testing Processes Reviewed (n=86)

Scoring Models and Instruments	All Departments (n = 86)		Large Departments (n = 29)		Medium Departments (n = 29)		Small Departments (n = 28)	
	#	% of total	#	% of total	#	% of total	#	% of total
<b><i>Single Component</i></b>	<b>67</b>	<b>78%</b>	<b>21</b>	<b>72%</b>	<b>23</b>	<b>79%</b>	<b>23</b>	<b>82%</b>
Written examination	8	9%	5	17%	1	3%	2	7%
Examination interview	36	42%	9	31%	12	41%	15	54%
Examination interview w/pre-interview exercise	13	15%	5	17%	7	24%	1	4%
Performance test	1	1%	0	--	1	3%	0	--
Supplemental application	0	--	0	--	0	--	0	--
E&E	9	10%	2	7%	2	7%	5	18%
<b><i>Multiple Weighted Components</i></b>	<b>13</b>	<b>15%</b>	<b>5</b>	<b>17%</b>	<b>3</b>	<b>10%</b>	<b>5</b>	<b>18%</b>
Written examination and examination interview	4	5%	3	10%	1	3%	0	--
Written examination and performance test	1	1%	1	3%	0	--	0	--
Performance test and examination interview	5	6%	0	--	1	3%	4	14%
Written examination, performance test, and examination interview	2	2%	1	3%	0	--	1	4%
Supplemental application and examination interview	0	--	0	--	0	--	0	--
Supplemental application and written examination	0	--	0	--	0	--	0	--
Supplemental application and performance test	1	1%	0	--	1	3%	0	--
<b><i>Multiple Components with Non-Weighted Components</i></b>	<b>6</b>	<b>7%</b>	<b>3</b>	<b>10%</b>	<b>3</b>	<b>10%</b>	<b>0</b>	<b>0%</b>
Examination interview (weighted 100%) and written examination	2	2%	0	--	2	7%	--	--
Examination interview (100%) and performance test	1	1%	0	--	1	3%	--	--
Performance test (100%) and written examination	1	1%	1	3%	0	--	--	--
Examination interview (100%) and supplemental application	1	1%	1	3%	0	--	--	--
Examination interview/performance test (100%) and written examination	1	1%	1	3%	0	--	--	--

## **Cost Comparison of Centralized and Decentralized Examination Processes**





### Office Technician Examination Development and Administration Costs

Cost Breakdown	Department	Departmental Costs	Estimated Costs for development of new examination	Estimated Costs for large candidate group and development of new examination	Estimated Duplicative Costs for large candidate group and development of new examination
Total Process	DMV (N = 54)	\$ 2,439.61	\$ 10,942.61	\$128,973.81	\$388,582.41
	CHP (N = 250)	\$ 5,759.19	\$ 14,491.19	\$129,934.80	
	CYA (N = 70)	\$ 3,412.70	\$ 9,895.70	\$129,673.80	
	SPB (N = 11544)	\$138,646.00	\$138,646.00	\$138,646.00	
By Applicant	DMV (N = 54)	\$45.18	\$202.64	\$11.17	\$33.66
	CHP (N = 250)	\$23.04	\$ 57.96	\$11.26	
	CYA (N = 70)	\$48.75	\$141.37	\$11.23	
	SPB (N = 11544)	\$12.01	\$ 12.01	\$12.01	
By Candidate	DMV (N = 41)	\$59.50	\$266.89	\$15.65	\$47.15
	CHP (N = 240)	\$24.00	\$ 60.38	\$15.76	
	CYA (N = 60)	\$56.88	\$164.93	\$15.73	
	SPB (N = 8242)	\$16.82	\$ 16.82	\$16.82	
By Eligible	DMV (N = 28)	\$87.13	\$390.81	\$26.61	\$80.19
	CHP (N = 113)	\$50.97	\$128.24	\$26.81	
	CYA (N = 35)	\$97.51	\$282.73	\$26.76	
	SPB (N = 4846)	\$28.61	\$28.61	\$28.61	

## **Standards for Decentralized Departmental Examination Programs**

## **Standards for Decentralized Departmental Examination Programs**

Departmental testing processes shall result in the job-related assessment of candidate qualifications. Departmental examination programs shall adhere to the following program and staffing standards:

### **Examination Program Standards**

The following program standards have been developed to ensure the quality and soundness of departmental testing processes. The standards provide specific direction to departments and agencies in the design and development of employment testing processes and individual selection instruments in conformance with existing legal and professional standards and principles as outlined by a myriad of legal and professional references, including the *Uniform Guidelines*, the Civil Rights Acts of 1964 and 1991, the Americans with Disabilities Act, the California Fair Employment and Housing Act, the *Principles for the Validation and Use of Personnel Selection Procedures*, and the *Standards for Educational and Psychological Testing*, as well as professional literature.

Civil service examination processes shall include the following:

1. The use of job-related scoring models for examination processes and the individual selection procedures comprising the examination processes
2. The use of soundly developed, job-related examination processes and individual selection procedures, developed based upon job analytic data, subject matter expert input/expertise, and in adherence to recognized legal and professional standards
3. The use of selection procedures appropriate to assess those KSAs identified as important for successful job performance and required-upon-entry to the job
4. The use of job-related scoring criteria to assess candidate qualifications and performance in the selection procedure/examination process
5. The use of job-related pass points
6. Documentation linking the content of the examination process and its individual selection procedures to the content requirements of the job classification for which the examination process is conducted



7. Documentation linking the scoring model of the examination process and its individual selection procedures to the content requirements of the job classification for which the examination process is conducted
8. The establishment of eligible lists (a) which are comprised of those candidates who demonstrate at least acceptable levels of competency to perform in the job classification for which the examination process is conducted and (b) which indicate in score-order those candidates who demonstrate greater levels of competency to perform in the job classification for which the examination process is conducted

### **Staffing Standards**

Each department and agency shall ensure that its examination program is managed by a technically competent unit manager and/or supervisor(s) possessing testing expertise to monitor and guide the work activities of Examination Unit staff in order to ensure the quality and soundness of the department's testing efforts, as follows:



**Examination  
Manager**

- Manage Examination Unit activities and projects
- Supervise Examination Unit supervisors and/or staff
- Serve as departmental testing expert and provide testing expertise in responding to selection-related questions and issues raised by examination staff and departmental staff, as well as other departments' examination units
- Ensure the quality and soundness of departmental testing activities
- Establish and implement performance expectations for Examination Unit staff
- Ensure that Examination Unit staff are fully trained and competent to conduct assigned testing activities
- Ensure departmental adherence to SPB policies and guidelines relating to testing activities
- Ensure departmental adherence to legal requirements pertaining to employment testing
- Ensure departmental adherence to recognized professional standards and principles pertaining to employment testing
- Promote the role of sound testing and selection in departmental hiring/staffing efforts
- Advocate sound testing and selection in departmental hiring/staffing efforts
- Stay abreast of current trends and practices in the assessment field as related specifically to employment testing



**Examination  
Supervisor**

- Assist in the management of/manage Examination Unit activities and projects
- Supervise Examination Unit staff
- Serve as departmental testing expert and provide testing expertise in responding to selection-related questions and issues raised by examination staff and departmental staff, as well as other departments' examination units
- Monitor the work of Examination Unit staff to ensure the quality and soundness of departmental testing activities
- Establish and/or implement performance expectations for Examination Unit staff
- Ensure that Examination Unit staff are fully trained and competent to conduct assigned testing activities
- Provide expert guidance and oversight to Examination Unit staff in the design, development, and administration of examination processes and individual selection procedures
- Monitor departmental testing activities to ensure adherence to SPB policies and guidelines relating to testing activities
- Monitor departmental testing activities to ensure adherence to legal requirements pertaining to employment testing
- Monitor departmental testing activities to ensure adherence to recognized professional standards and principles pertaining to employment testing
- Promote the role of sound testing and selection in departmental hiring/staffing efforts
- Advocate sound testing and selection in departmental hiring/staffing efforts
- Stay abreast of current trends and practices in the assessment field as related specifically to employment testing



In addition, each department and agency shall ensure that its examination program is staffed by a technically proficient analyst staff possessing assessment training and expertise consistent with either journey-level or entry-level competency to adequately perform the duties of an examination analyst with sufficient technical supervisory oversight, as follows:

**Examination Analyst**

(journey-level performance)

- Conduct job analysis studies of varying degrees of complexity utilizing recognized, legally defensible job analysis methodologies
- Design and develop job-related, content-valid, legally defensible examination processes and individual selection procedures
- Administer and/or oversee the administration of a variety of selection procedures, ensuring accurate and standardized administration for all candidates
- Document the content validity of examination processes and individual selection procedures developed and/or administered as part of departmental testing activities
- Interpret item analyses and examination statistics to determine instrument soundness and the extent to which instrument performance could be improved
- Establish job-related pass points for a variety of selection procedures using examination statistics, empirical data, and subject matter expert input/expertise
- Provide testing expertise in responding to selection-related departmental questions and issues
- Stay abreast of current trends and practices in the assessment field as related specifically to employment testing

